

**GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE**  
**ANTITUMOR ANTIBIOTIC C-1027**

**SEQUENCE LISTING**

SEQ ID No. 1. C-1027 gene cluster DNA sequence from 1 to 42,000, ORF-(-7) to ORF-26

1 GTCGACTCTAGAGGATCCCGGGTGCAGGAGTAGGGGTTACGGACGAAGGAGGGGTGCCCCG  
-----+-----+-----+-----+-----+ 60  
-7-\* CAGCTGAGATCTCCTAGGGCCACGCCTCATCCCCAATGCCTGCTTCCTCCCCACGGGCC  
\* L I G P A S Y P N R V F S P H G -  
  
61 CGACGCCTGCGGCGAAGGGCGGTTCTTCTGAGTTCGAGGCCGGTGGCGAGGACGACGTGGT  
-----+-----+-----+-----+-----+ 120  
-7 GCTGCGGACGCGCTTCCCAGCAAGGAATCAAGTCCCGGCCACCGCTCCTGCTGCACCA  
A V G A A F P P E K L E L G T A L V V H -  
  
121 CCGCGTCGAGGATCTGCGTGTGCGGGAGCGGCCAGGGCGCAGCCCCCTCGGTGAGGTACG  
-----+-----+-----+-----+-----+ 180  
-7 GGCGCAGCTCCTAGACGCACAGCCCCCTCGCGGGTCCCGCGTCGCGGGAGCCAGTCCATGC  
D A D L I Q T D P L P G P R L G E T L Y -  
  
181 GGGTGAGGCCCTGACGGTCACCTCGAAGCAGCGGTCTGTTGGACCGGGCGTCGAGCGCCT  
-----+-----+-----+-----+-----+ 240  
-7 CCCACTCCGGGGACTGCCAGTGGAGCTTCGTCGCCAGCACCTGGCCCCGAGCTCGCGGA  
P T L G R V T V E F C R D H S R A D L A -  
  
241 CCCCCTCCGCTTCCACAAGGACGACCGCGGGACAGGACTCCCGTGCGGCCTCGACCAGTC  
-----+-----+-----+-----+-----+ 300  
-7 GGGGCAGGCGAAGGTGTTCTGCTGCGGCCCTGCTCTGAGGGCAGCGCGGAGCTGGTCAG  
E G D A E V L V V G P C S E R A A E V L -  
  
301 GGGCGTCGAGGTAGTCTTGAAGATGCGGCGGGGGCGGGGCCCTGTTCTGGTGAAGTTC  
-----+-----+-----+-----+-----+ 360  
-7 CCCGCAGCTCCATCAGGACCTTCTACGCCGCCCGCGCCCCGCGGACCAAGCCACTTGAAGG  
R A D L Y D Q F I R R P A P G Q E T F K -  
  
361 ACGAAGCCCAGCGCCGGGGCCAGTCGCGCCGGTCCGCTCCTGGTTGGCCCAGTTGATGA  
-----+-----+-----+-----+-----+ 420  
-7 TGCTTCGGGTGCGGCCCCGGTCAGCGCGGCCAGCGGAGGACCAACCGGGTCAACTACT  
W S A W R R P W D R R D A E Q N A W N I -  
  
421 AGTCGAGCACGTCTCGCGGAACACCGACATCCTGCGGCCTGGATATTGAAGACGTGGT  
-----+-----+-----+-----+-----+ 480  
-7 TCAGCTCGTGCAGGAGCGCCTTGTGGCTGTAGGACGCCCGGACCTATAACTTCTGCACCA  
F D L V D E R F V S M R G A Q I N F V H -  
  
481 CCCAGGGGTTGCCGTACGGTGATAGGCGACGCCGGCCGAGCGGTAGCGGGCGCGCCGCT  
-----+-----+-----+-----+-----+ 540  
-7 GGGTCCCCAACGGCAGTGCCACTATCCGCTGCGGGCCGCTCGCCATCGCCGCGCGGCGA  
D W P N G D R H Y A V G A S R Y A A R R -  
  
541 CCAGGAGGACGACTTCCAGCGGTCTTCTCGCGAAATGAAGCAGGCGTATCGCGGTGCGCCG  
-----+-----+-----+-----+-----+ 600  
-7 GGTCTCTGCTGAAGGTGCGCCAGAAGAGCGCTTTACTTCGTCCGCATAGCGCCAGCGGC  
E L L V V E L P R R A F H L L R I A T A -  
  
601 TGCCTGCCAGGCCCCGCCCTACGACCAGCACCTGGGGCGCGCACCCGTCATGCCCCATGA  
-----+-----+-----+-----+-----+ 660  
-7-< ACGGACGGTCCGGGCGGGGATGCTGGTCTGGGACCCCGCGGTGGGCAGTACGGGTACT  
T G A L G A G V V L V R P R A G T M G M -

AGCCTCCCCCGCTGACTCAGGGCG<sub>g</sub>CGCGTCGCGCGCTCCCGTCGGTGTCTCGCTGACT  
 661 -----+-----+-----+-----+-----+-----+ 720  
 TCGGAGGGGGCGACTGAGTCCCGC<sub>c</sub>GCGCAGCGCGGAGGGCAGCCACAGGAGCGACTGA  
  
 GGAAGTTCCTGACCTGGCGTCAACTCCACTGATCCGTAAGGGGATCGCGGGAGTGGATA  
 721 -----+-----+-----+-----+-----+-----+ 780  
 CCTTCAAGGGACTGGACCGCAGTTGAGGTGACTAGGCATTCCCCTAGCGCCCTCACCTAT  
  
 CGGGTCAGGTTCGTGCACGATCGTGGCACCAGACAGATCACCACGTGATAGGCACTCGTG  
 781 -----+-----+-----+-----+-----+-----+ 840  
 GCCCAGTCCAGCACGTGCTAGCACCGTGGTCTGTCTAGTGGTGCAGCTATCCGTGAGCAC  
  
 AGCCGCGCCCGGGGCTCGACGGGGCGGGGCACCGGCAGGGGCGGCCGCGTGATCAGCCGG  
 841 -----+-----+-----+-----+-----+-----+ 900  
 TCGGCGCGGGCCCCGAGCTGCCCCGCCCCGTGGCCGTCCCCGCGGCGCACTAGTCGGCC  
  
 AGCCTGTCCGGGGGCGTGCGTGC GG GCGTCAGCTGTTCGATGTTCGGAACGCCAGGGACG  
 901 -----+-----+-----+-----+-----+-----+ 960  
 TCGGACAGGCCCCCGCACGCACGCCCCGAGTCGACAGCTACAGCCCTTGCGGTCCCTGC  
 -6-\* \* S D I D P V G P V -  
  
 TCGATCTCGGTGCGGGCGTAGTGGTTGAAGTAGTTGGTGTAGAGGTTACAGGCCACGTGG  
 961 -----+-----+-----+-----+-----+-----+ 1020  
 AGCTAGAGCCACGCCCCGCATCACCAACTTCATCAACCACATCTCCAAGTGCCGCTGCACC  
 -6 D I E T R A Y H N F Y N T Y L N V A V H -  
  
 ACGAAGACCTCGGCGAGCTCGGTGTCCGTCCATCCCTGTGCCACGGCCGCGTTCCACGAG  
 1021 -----+-----+-----+-----+-----+-----+ 1080  
 TGCTTCTGGAGCCGCTCGAGCCACAGGCAGGTAGGGACACGGTGCCGGCGCAAGGTGCTC  
 -6 V F V E A L E T D T W G Q A V A A N W S -  
  
 GCGTCAGACGCCTCGCCCACTTCGCCGGCGATCTCCCTGGCCACCTGGACCAGTGCTTCG  
 1081 -----+-----+-----+-----+-----+-----+ 1140  
 CGCAGTCTGCGGAGCGGGTGAAGCGGCCGCTAGAGGGACCGGTGGACCTGGTCACGAAGC  
 -6 A D S A E G V E G A I E R A V Q V L A E -  
  
 AGCTTCACGTTCGTGCGCGGGCGTCCCCCGGCGAATCGCCACGGTCTCCTCCAGCGTGAAA  
 1141 -----+-----+-----+-----+-----+-----+ 1200  
 TCGAAGTGCAGCAGCGGCCCGCAGGGGGCCGCTTAGCGGTGCCAGAGGAGTTCGCACTTT  
 -6 L K V D D G P T G R R I A V T E E L T F -  
  
 CCCGCGACCTTCGCCGACACCGTGTGCGCCGCTGGCAGTACGCGCACGCGTCGACCGCG  
 1201 -----+-----+-----+-----+-----+-----+ 1260  
 GGGCGCTGGAAGCGGCTGTGGCACACGCGGCGGACCGTCATGCGCGTGCGCAGCTGGCGC  
 -6 G A V K A S V T H A A Q C Y A C A D V A -  
  
 CCCACGGCGAGGGCGATCGCCTCGCGTGTGCGGGCGTGAACGTTCCATGTTTCGGCGACG  
 1261 -----+-----+-----+-----+-----+-----+ 1320  
 GGGTGCCGCTCCCGCTAGCGGAGCGCACACGCCCCGAGCTTGCAAGGTACAAGCCGCTGC  
 -6 G V A L A I A E R T R A D F T G H E A V -  
  
 GCTCCGGTGATCGCGGCGTAGGTTTCCAGGACCACGGGGGAATGGGCCATTCCCCCGTGG  
 1321 -----+-----+-----+-----+-----+-----+ 1380  
 CGAGGCCACTAGCGCCGCATCAAAGGTCCTGGTGCCCCCTTACCCGGTAAGGGGGCACC  
 -6 A G T I A A Y T E L V V P S H A M G G H -  
  
 ATGTTGAGCACTCGCCCGAACCGCTTCTCCAGTCGGCGCAGGATGTCTCCGCCGGCTGCG  
 1381 -----+-----+-----+-----+-----+-----+ 1440  
 TACAACTCGTGAGCGGGCTTGGCGAAGAGGTGAGCCGCGTCTACAGAGGCGGCCGACGC  
 -6 I N L V R G F R K E L R R L I D G G A A -  
  
 GGTGCGGTGTTCGATGGTGTGGACGGGAATCCGCGGCATGGGAATGCCTCTCCTCGTAGTG  
 1441 -----+-----+-----+-----+-----+-----+ 1500  
 CCACGCCACAGCTACCACACCAGCCCTTAGGCGCCGTACCCTTACGGAGAGGAGCATCAC  
 -6-< P A T D I T H V P I R P M -

054304500

```

ATGGGAGTTCTCGTCCCTCCAGTCTGCCAAGCACCTCCCCCGGTGAGCTGTCCCGGCC
1501 -----+-----+-----+-----+-----+-----+ 1560
TACCCTCAAGGAGCAGGGAGGTGACACGGGTTCTGAGAGGGGGCCACTCGACAGGGCCGG

GCCCTCCGGCCCCCTTCTAGGCAGGTGCCCCGGTGGTGCGGCCCCAGGACGTACCTCGCC
1561 -----+-----+-----+-----+-----+-----+ 1620
CGGGAGGCCGGGAAGATCCGTCCAGCGGGCCACCACGCCGGGGTCTGCACTGGAGCGG

GCACCACCGGGAGCCCCGAGGGGCGAGGTGAGAGGCCGAGCACCTCTCGGCCAGGGCGG
1621 -----+-----+-----+-----+-----+-----+ 1680
CGTGGTGGCCCTCGGGGCTCCCCGCTCCAGTCTCCGGCTCGTGGAGGAGCCGGTCCCGCC
-5- * * L G L V E E A L A -

TGCCCCGAACACGGGCCTCGATCTTGGCGAAGGCCAGGTGCGGTGTGGTGGAGGTGTCTG
1681 -----+-----+-----+-----+-----+-----+ 1740
ACGGGGCTTGTGCCCGGAGCTAGAACCCTTCCGGTCCAGCGCACACCACCTCCACAGCA
-5 T G R V R A E I K A F A L D R T T S T D -

CGCGAACGGGGAGAAGCCGAGTCTGCGAGGTTCCAGTTGCTCGACGGGGATGTAGC
1741 -----+-----+-----+-----+-----+-----+ 1800
GCCGCTTGCCCCCTTCTCGGCGTCAAGAGCTCAACGAGCTGCCCTACATCG
-5 D A F P S F G C D D C T G L Q E V P I Y -

GGGCGGCGAGCAGGATGCGGTGCGGTACCTGCTCGGGGGTCTCGACCACTGGGTGATCG
1801 -----+-----+-----+-----+-----+-----+ 1860
CCCGCGCTCGTCTACGCCAGCGCATGGACGAGCCCCCAGAGCTGGTGACCCAGCTAGC
-5 R A A L L I R D R V Q E P T E V V P D I -

GGTCGGTCACCCCGAGGAAGACGCGGGCGGCAGGGGGCAGGTGGTCACGGACGATGCTCA
1861 -----+-----+-----+-----+-----+-----+ 1920
CCAGCCAGTGGGGCTCCTTCTGCGCCCGCTCCCCCGTCCACCACTGCTGCTACGAGT
-5 P D T V G L F V R A A P P L H D R V I S -

GGACCCGCTCGGGGTCCGCTTCCGCCGCCAGTTCGAGATAGAAGTTGCCCGCCTTGAGCT
1921 -----+-----+-----+-----+-----+-----+ 1980
CCTGGGCGAGCCCCAGGCGAAGCGGCCGGTCAAGCTCTATCTTCAACGGGCGGAACCTGA
-5 L V R E P D A E G A L E L Y F N G A K L -

GGAAGAGCTTGGGCGAGTTCGGCGTAGTCGATGTCGAGGCTGTGCGTGGAGTCTGGT
1981 -----+-----+-----+-----+-----+-----+ 2040
CCTTCTCGAACCCGTCGTCAAGCCGCATCAGCTACAGCTCCGACACGCACCTCAGGACCA
-5 Q F L K P L L E A Y D I D L S H T S D Q -

CGCCGCCGGGGCAGGTGTGTACGCCGATGCGGGCGGTTTCTCGGCGCTGAAGCGCCCCA
2041 -----+-----+-----+-----+-----+-----+ 2100
GCGGCGGCCCCGTCCACACATGCGGCTACGCCCGCAAAGGAGCCGCGACTTCGCGGGGT
-5 D G G P C T H V G I R A T E E A S F R G -

GGACTTCGTTGTTGAGGGCGATGAAGTCGTCGAGGACGCCCGCTGGGGTTCGAGCTTGA
2101 -----+-----+-----+-----+-----+-----+ 2160
CCTGAAGCAACAACCTCCGCTACTTCAGCAGCTCCTGCGGCGGCGACCCAGCTCGAACT
-5 L V E N N L A I F D D L V G G S P D L K -

GGGACAGCCGCCCTCGGTGAAGTCGAGCTGGACCACGTGTGCCCCCGCTCCAGGCAGC
2161 -----+-----+-----+-----+-----+-----+ 2220
CCCTGTGCGGGGAGCCACTTCAGCTCGACCTGGTGCACACGGGGGCGCAGGTCCGTGCG
-5 L S L R G E T F D L Q V V H A G A D L C -

CTCGGATGTCGGCTTCGGCCTCGTCCGCGAGGTGCGCGCAGGAAGTCTCGCGGGGGTAGC
2221 -----+-----+-----+-----+-----+-----+ 2280
GAGCCTACAGCCGAAGCCGGAGCAGCCGCTCCAGCGCGTCTTACGAGCGCCCCCATCG
-5 G R I D A E A E D A L D R L F Q E R P Y -

CCTCGATGGGAGTGGCGGGGTAGAGGAGGTGAGGGCGGAGGGTGCATGACCGCCTGCT
2281 -----+-----+-----+-----+-----+-----+ 2340

```







-3           A K E G G A R K W R Y G D W S Y W G L G -  
 CTCCACGTGCCGGAAC TGGT CACGGT AGAGAC GGCAGG AAGAGCTTGGCGTTGCCGCGGTC  
 4741 -----+-----+-----+-----+-----+-----+-----+-----+ 4800  
 GAGGTGCACGGCCTTGACCA GTGCCATCTCTGCCTGCTTCTCGAACCGCAACGGCGCCAG  
 -3           E V H R F Q D R Y L R V F L K A N G R D -  
 GGTCAGGCTGGCGGGAATCTCGCCCGCCTCCAGGCGGTTCGCGGCGACGGGGGCTCGGG  
 4801 -----+-----+-----+-----+-----+-----+-----+ 4860  
 CCAGTCCGACCGCCCTTAGAGCGGGCGGAGGGTCCGCCAGCGCCGCTGCCCCCGGAGCCC  
 -3           T L S A P I E G A E W A T A A V P A E P -  
 AGCGGCCTGGACAGGGAGGAGCGGCGCTGGGGCCGGGGTGGTTTCGAGGGCCAGCATCTG  
 4861 -----+-----+-----+-----+-----+-----+-----+ 4920  
 TCGCCGGACCTGTCCCTCCTCGCCGCGACCCCGGCCACCAAAGCTCCCGGTCTGAGAC  
 -3           A A Q V P L L P A P A P T T E L A L M Q -  
 CTGAGCGGCGGCAGTTGCGTCAAAGCGAGGGCCCTCGGCGCTGCTGCTCATGGACGTCTCT  
 4921 -----+-----+-----+-----+-----+-----+-----+ 4980  
 GACTCGCCCGCTCAACGCAGTTTCGCTCCCGGGAGCCGCGACGACGAGTACCTGCAGGA  
 -3-<           Q A A A T A D F R P G E A S S S M -  
 TCGAGATGGAGCGGTTCGGGCGGTCCCGCTGCGGGAACGGCATGAATGATCTTCCCGGTG  
 4981 -----+-----+-----+-----+-----+-----+-----+ 5040  
 AGCTCTACCTCGCCAGCCCGCCAGGGGCGACGCCCTTGCCGTACTTACTAGAAGGGCCAC  
 CGGACAGAGTGCCAGGGGCGAGCGCATGTGCGGGGGGACAACGGCCCGTTTCGGACGAGGG  
 5041 -----+-----+-----+-----+-----+-----+-----+ 5100  
 GCCTGTCTCACGGTCCCCGTGCGTACACGCCCCCTGTTGCCGGGCAAAGCTGCTCCC  
 CCGGCCGACGGGGGAAGCAGGGGCGGCAACCGGGTGGCGGGGCGGCGTGAGCGAGGGC  
 5101 -----+-----+-----+-----+-----+-----+-----+ 5160  
 GGCCGGCTGCCCCCTTCGTCCCCGCGCTTGGCCCACCGCCCCGCGCACTCGCTCCCC  
 ACGAGCGGCCCGGTACGGGGGGAAGGGCTCGTCTCTCCGTGGGGGCGGCACGTTGTGGTCC  
 5161 -----+-----+-----+-----+-----+-----+-----+ 5220  
 TGCTCGCCGGGCCATGCCCCCTTCCCGAGCAGAGAGGCACCCCGCGTGCAACACCAGG  
 TCGTCCGTGAGCTTGGCTTTCAGCCTCCTGACCCCCAATAAGGCGAAAGCTGCTG  
 5221 -----+-----+-----+-----+-----+-----+-----+ 5280  
 AGCAGGCAGTCGAACGCAGACCGAAGTCGGAGGACTGGGGGTATTCCGCTTTCGACGAC  
 GTCAAGCATCTTTTCGTGACACTCGGCGAGGGACTGAAGGGACTGTCTTTCGGAATGAGTG  
 5281 -----+-----+-----+-----+-----+-----+-----+ 5340  
 CAGTTCGTAGAAAGCACTGTGAGCCGCTCCCTGACTTCCCTGACAGAAAGCCTTACTCAC  
 TAGGGGGTTGTGCGGTGGGGACCGCGCCTCGACTCCCCGGCGGACGGGATCTGTTGCGTTC  
 5341 -----+-----+-----+-----+-----+-----+-----+ 5400  
 ATCCCCCAACAGCCACCCCTGGCGCGGAGCTGAGGGGCGCCTGCCCTAGACAAGCCAG  
 GGTCCCTTGGGTCCCTCCCCGGATCGCGGCAGGGACCCAAGGGGGCGGTGCGGCGGGCGG  
 5401 -----+-----+-----+-----+-----+-----+-----+ 5460  
 CCAGGGAACCCAGGGAGGGGCCTAGCGCCGTCCCTGGGTCCCCCGCCACGCCGCCCGCC  
 TCGGTGAGGGGCCCCGGTGGAGGGACTGAGGGTCTGTATGGAGCGATAAGAGGGTCTGAA  
 5461 -----+-----+-----+-----+-----+-----+-----+ 5520  
 AGCCACTCCCCGGGGCCACCTCCCTGACTCCCGACATACCTCGCTATTCTCCAGACTT  
 GGGGCGGAGAGAGTTTCGGTCCCTGCGTTGAGTCCCTGGTCATCACCGCAGGTGAGAGGG  
 5521 -----+-----+-----+-----+-----+-----+-----+ 5580  
 CCCCCGCTCTCTCAAAGCCAGGGACGCAACTCAGGGACCAGTAGTGGCGTCCAGTCTCCC  
 GTTTTGAGGGGTGAAAAAGGGACTGAAGGGACTCAACTTCCCCATTATGAGCTGAGTAGA  
 5581 -----+-----+-----+-----+-----+-----+-----+ 5640  
 CAAAACTCCCCACTTTTTCCCTGACTTCCCTGAGTTGAAGGGTAATACTCGACTCATCT







```

TCAACTTCCGCAGGCTGGCCAGTGGCCGCCGCGAGCCGCGGACGAGCCGCTGAAGCGCC
-2      L K A S D R S P A A L G A L L G D F A A -

CCGTCAACGAGGTCATCACCGCAGCGATGGACGATGTCCTGCGCAGTGGAGCGGACCCCG
7321    +-----+-----+-----+-----+ 7380
GGCAGTTGCTCCAGTAGTGGCGTCTGCTACCTGCTACAGGACGCGTCACCTCGCCTGGGGC
-2      V N E V I T A A M D D V L R S G A D P A -

CGAAGGCCTTCGCGGAAGCCGGCGTGGCCGCCAGCAACTGCTCGATGCCTACAACGCCC
7381    +-----+-----+-----+-----+ 7440
GCTTCCGGAAGCGGCTTCGGCCGACCGGCGGGTCTGACGAGCTACGGATGTTGCGGG
-2      K A F A E A G V A A Q Q L L D A Y N A R -

GGAACCGCTCCGGATCCGGGACCCCTCCGCCGTCTGAGATCCGGTACCGGGGCACAGGG
7441    +-----+-----+-----+-----+ 7500
CCTTGGCGAGGCCTAGGCCCTGGGGGAGGCGGCAGACTCTAGGCCATGGCCCCGTGTCCC
-2-*    N R S G S G T P S A V * -

GCGCCGCCCGCCGCTTTCCCGGCGGGGCACTGGCCGGGGGACATGCTCTCCCGCCCCCGG
7501    +-----+-----+-----+-----+ 7560
CGCGGCGGCGGGCGAAAGGGCCGCCCGTGACCGGCCCCCTGTACGAGAGGGCGGGGGCC

CAGGACGTAGGGTCAACCCGCTGCGCCTTCAGGTGGCGGCGCAGATACTCACCGGTCAG
7561    +-----+-----+-----+-----+ 7620
GTCCTGCATCCCAGTTGGGCGGACGCGGAAGTCCACCGCCGCGTCTATGAGTGGCCAGTC
-1-*    * G A Q A K L H R R L Y E G T L -

GGAGGAATCCGCGGCGAGCAGGTCTTTCGGTGTGCCGGTGAAGACGATCTCGCCGCCCTC
7621    +-----+-----+-----+-----+ 7680
CCTCCTTAGGCGCCGCTCGTCCAGGAAGCCACACGGCCACTTCTGCTAGAGCGGCGGGAG
-1      S S D A A L L D K P T G T F V I E G G E -

CCGTCCCCCGTCGGGACCCAGGTTCGATGATCCAGTTCGGCCTGCTGCACCACATCGAGGTT
7681    +-----+-----+-----+-----+ 7740
GGCAGGGGGCAGCCCTGGGTCCAGCTACTAGGTACGCCGACGACGTGGTGTAGCTCCAA
-1      R G G D P G L D I I W D A Q Q V V D L N -

GTGCTCGATGACCACGACGGTGTTCGCCGCCCTCGACGAGCCCGTCCAGGAGCTTCAGCAG
7741    +-----+-----+-----+-----+ 7800
CACGAGCTACTGGTGTGCCACAAGGGCCGGAGCTGCTCGGGCAGGTCTCGAAGTCGTC
-1      H E I V V V T N G A E V L G D L L K L L -

GGTGTCAACGTCCGACATGTGCAGCCCGGTGGTGGGCTCGTCCAGGACATAGACCGTGCC
7801    +-----+-----+-----+-----+ 7860
CCACAGTTGCAGGCTGTACACGTCCGGGCCACCAACCGAGCAGGTCTGTATCTGGCACGG
-1      T D V D S M H L G T T P E D L V Y V T G -

CGTGCGGTGCAGCTGGTTCGGCAAGTTTGATCCGCTGCAGTTACCGCCGGAGAGGCTGGA
7861    +-----+-----+-----+-----+ 7920
GCACGCCACGTTCGACCAGCCGTTCAAAGTAGGCGACGTCAAGTGGCGGCTCTCCGACCT
-1      T R H L Q D A L K I R Q L E G G S L S S -

AAGCGGCTGGCCCAGGCTGAGGTACCCAAGACCGACGTTCGACGAGAGCGCGAGTTTCGG
7921    +-----+-----+-----+-----+ 7980
TTCGCCGACCGGGTCCGACTCCATGGGTCTGGCTGCAGCTGCTCTCGCGCGTCAAAGCC
-1      L P Q G L S L Y G L G V D V L A R L K P -

CAGCAGGGCCTTCTCGGTGAAGAACTCGACGGCCTCGTCGGCGGGCAGCTCCAGGACGTC
7981    +-----+-----+-----+-----+ 8040
GTCGTCCCGGAAGAGCCACTTCTTGAGCTGCCGGAGCAGCCGCCCGTCGAGGTCTTCAG
-1      L L A K E T F F E V A E D A P L E L V D -

CGCGATCGACTTCCCGCGAAGCTGGTGTCTCCAGGACCTCGGGCTTGAAGCGGCGCCCCCTC
8041    +-----+-----+-----+-----+ 8100
GCGCTAGCTGAAGGGCGCTTCGACCACGAGGTCTGGAGCCCGAAGTTCGCCGCGGGGAG
-1      A I S K G R L Q H E L V E P K F R R G E -

```

```

ACAGACACCGCAGTGGTGGTCACCGGATCCATGAAGGCCAGCTCGGTGATGATGACCCC
8101 -----+-----+-----+-----+-----+-----+-----+-----+ 8160
TGTCTGTGGCGTCACGCACCGAGTGGCCTAGGTACTTCCGGTCGAGCCACTACTACTGGGG
-1 C V G C H T T V P D M F A L E T I I V G -

GCGGCCCTGGCACTCCTCGCACGACCCCTTGGAGTTGAAGCTGAACAGCGAGGCGTTCGC
8161 -----+-----+-----+-----+-----+-----+-----+-----+ 8220
CGCCGGGACCGTGAGGAGCGTGCTGGGGAACCTCAACTTCGACTTGTCTCGCTCCGCAAGCG
-1 R G Q C E E C S G K S N F S F L S A N A -

GCCGGTCTCCTTCGCGAACAGCTTGCAGCAGCGGGTCCATCAGGCCGAGGTAGGAGACCGG
8221 -----+-----+-----+-----+-----+-----+-----+-----+ 8280
CGGCCAGAGGAAGCGCTTGTGCAACGCGTCGCCCAGGTAGTCCGGCTCCATCCTCTGGCC
-1 G T E K A F L K R L P D M L G L Y S V P -

TGTGGAGCGCGACGAGGCGGCGATCGCGGACTGGTTCGACAAAGACCGCGTCGGGGTGC
8281 -----+-----+-----+-----+-----+-----+-----+-----+ 8340
ACACCTCGCGCTGCTCCGCCGCTAGCGCCTGACCAGCTGTTTCTGGCGCAGCCCCACGCG
-1 T S R S S A A I A S Q D V F V A D P H A -

CTCCATGAATGCCCCGGAGATCAGGCTGCTCTTGCCGGAACCCGCCACCCCGGTACCCGC
8341 -----+-----+-----+-----+-----+-----+-----+-----+ 8400
GAGGTACTTACGGGGCCTCTAGTCCGACGAGAACGGCCTTGGGCGGTGGGGCCAGTGGCG
-1 E M F A G S I L S S K G S G A V G T V A -

GGTCAGCACACCGGTGGGCACGGCCACGGAGACCTGCTTCAGGTTGTGGAGATCCGCGTT
8401 -----+-----+-----+-----+-----+-----+-----+-----+ 8460
CCAGTCGTGTGGCCACCCGTGCCGGTGCCTCTGGACGAAGTCCAACACCTCTAGGCGCAA
-1 T L V G T P V A V S V Q K L N H L D A N -

CTCCACGGTCAGCTCCCCCGTGGGCGGGCGGACCTCCTCCTTCACGCGGGCCCCCGCGG
8461 -----+-----+-----+-----+-----+-----+-----+-----+ 8520
GAGGTGCCAGTCGAGGGGGCACCCGCCCGCCTGGAGGAGGAAGTGCGCCCGGGGGCGGC
-1 E V T L E G T P P R V E E K V R A G R R -

CAGAGCCTCCCCGGTCCGGGTCTTCGCCTTCCGCAGCTTCGCGAAGGACCCCTCGAACAC
8521 -----+-----+-----+-----+-----+-----+-----+-----+ 8580
GTCTCGGAGGGGCCAGGCCCAGAAGCGGAAGGCGTGAAGCGCTTCTTGGGAGCTTGTG
-1 L A E G T R T K A K R L K A F S G E F V -

GATCTCGCCCCCGTGCACTCCCGCCCCGGGACCGACATCGACGATGTGGTTCGGCGATCTC
8581 -----+-----+-----+-----+-----+-----+-----+-----+ 8640
CTAGAGCGGGGGCACGTGAGGGCGGGGCCCTGGCTGTAGCTGCTACACCAGCCGCTAGAG
-1 I E G G H V G A G P G V D V I H D A I E -

GATCACATcGGGGTCGTGCTCGACGACCAGCACGGTGTTCCTTGTTCGCGCAGCGCGCG
8641 -----+-----+-----+-----+-----+-----+-----+-----+ 8700
CTAGTGTAgCCCCAGCACGAGCTGCTGGTCTGCCACAAGGGGAACAGCGCTCGCGCGC
-1 I V D P D H E V V L V T N G K D R L A R -

CAGCAGGTCGTTGAGCCGCCCCACGTGCGCGGGGTGCAGGCCGATGCTGGGCTCGTCGAA
8701 -----+-----+-----+-----+-----+-----+-----+-----+ 8760
GTCGTCCAGCAACTCGGCGGGGTGCAGCGCGCCACGTCCGGCTACGACCCGAGCAGCTT
-1 L L D N L R G V D R P H L G I S P E D F -

GATGTACGTGAGCCCGGCCAGACCACTGCCGAGGTGGCGCACCATCTTCAGCCGCTGCCC
8761 -----+-----+-----+-----+-----+-----+-----+-----+ 8820
CTACATGCACTCGGGCCGGTCTGGTGACGGCTCCACCGCGTGGTAGAAGTCGGCGACGGG
-1 I Y T L G A L G S G L H R V M K L R Q G -

CTCGCCCCCGAGAGGTTCGGCCGTGGGCCTGTCCAGGGTCAGGTAGCCGAGCCCGATGGA
8821 -----+-----+-----+-----+-----+-----+-----+-----+ 8880
GAGCGGGGGGCTCTCCAGCCGGCACCCGGACAGGTCCAGTCCATCGGCTCGGGCTACCT
-1 E G G S L D A T P R D L T L Y G L G I S -

```

CACGATCCGCTCCAGGGCCGTGCGCGCGGCTTTCGCGAGAGGGGCAGCGGCCGGCTCCGT  
8881 -----+-----+-----+-----+-----+-----+-----+ 8940  
GTGCTAGGCGAGGTCCCGGCACGCGCGCCGAAAGCGCTCTCCCGTCGCCGGCCGAGGCA  
-1 V I R E L A T R A A K A L P A A A P E T -  
  
GACGCCGGCGAGCACCTCCGTGAGGTCGCGGACCTCCATGCTCGAGTAGTCGGCGATGTT  
8941 -----+-----+-----+-----+-----+-----+-----+ 9000  
CTGCGGCCGCTCGTGGAGGCACTCCAGCGCCTGGAGGTACGAGCTCATCAGCCGCTACAA  
-1 V G A L V E T L D R V E M S S Y D A I N -  
  
CTTGCCGTCGATCCGGACGTCGAGCGCGGCGGCGTTGAGCCGCGCGCCCCGGCAGGAGGG  
9001 -----+-----+-----+-----+-----+-----+-----+ 9060  
GAACGGCAGCTAGGCCTGCAGCTCGCGCCGCCGCAACTCGGCGCGCGGGGCCGTCTCTCC  
-1 K G D I R V D L A A A N L R A G R C S P -  
  
ACAGACTCCGTCgGTGACGAAACGTTTCGATGACCTCGCGCTTGCGGTTCgTCAGCGCGCT  
9061 -----+-----+-----+-----+-----+-----+-----+ 9120  
TGTCTGAGGCAGcCACTGCTTTGCAAGCTACTGGAGCGCGAACGCCAGCgAGTCGCGCGA  
-1 C V G D T V F R E I V E R K R D S L A S -  
  
GAGGTCGCGCTTGAGGTTGAgCCGCTCGAACCGGTTCGGCcAACCCCTCGTAGTTCTGTCG  
9121 -----+-----+-----+-----+-----+-----+-----+ 9180  
CTCCAGCGCGAACTCCAACtcGGCGAGCTTGGCCAGCCGgTTGGGGAGCATCAAGCAGAC  
-1 L D R K L N L R E F R D A L G E Y N T Q -  
  
GAACTCGGTGCTCTTGGTCTTCAGCGTcACCTTCCCGCCGGTTCgCGCGCAGCAGCGTGTC  
9181 -----+-----+-----+-----+-----+-----+-----+ 9240  
CTTGAGCCACGAGAACCAGAAGTCGAgTGAAGGGCGGCCACgGCGCGTCGTCGCACAG  
-1 F E T S K T K L T V K G G T G R L L T D -  
  
CAGCTCCTCGGCGCTGTACTCGGCGATCGGCTTGGCCGGATCCAGACGGCCGGACTTCGC  
9241 -----+-----+-----+-----+-----+-----+-----+ 9300  
GTCGAGGAGCCGCGACATGAGCCGCTAGCCGAACCGGCCTAGGTCTGCCGGCCTGAAGCG  
-1 L E E A S Y E A I P K A P D L R G S K A -  
  
CCAGATCTGCCAGTCCGGGCTACCCACCTTGTACTCGGGGAAAAAGACCGCCCCGTCGTC  
9301 -----+-----+-----+-----+-----+-----+-----+ 9360  
GGTCTAGACGGTCAGGCCCCGATGGGTGGAACATGAGCCCTTTTCCTGGCGGGGCAGCAG  
-1 W I Q W D P S G V K Y E P F L V A G D D -  
  
CAGGGACTTCGAGCGGTCCAGCATCTTGTCCAGGTCGAGGGCGATGCTCTGGCCGAGACC  
9361 -----+-----+-----+-----+-----+-----+-----+ 9420  
GTCCCTGAAGCTCGCCAGGTCGTAGAACAGGTCCAGCTCCCGCTACGAGACCGGCTCTGG  
-1 L S K S R D L M K D L D L A I S Q G L G -  
  
GTCGCAGTCCGGGCACATGCCCTGGGGGTCGTTGAACGAGAACCGGAGACGCCGAGCGA  
9421 -----+-----+-----+-----+-----+-----+-----+ 9480  
CAGCGTCAGGCCCCGTGTACGGGACCCCCAGCAACTTGCTCTTGCGCCTCTGCGGCTCGCT  
-1 D C D P C M G Q P D N F S F A S V G L S -  
  
GGACGGCCCCGTCGTCCTTCGTCGTCGCGAACCGTGCGAACAGGGCCCGGATCATCGGCTG  
9481 -----+-----+-----+-----+-----+-----+-----+ 9540  
CCTGCCGGGCAGCAGGAAGCAGCACGGCTTGGCACGCTTGTCCCGGCCCTAGTAGCCGAC  
-1 S P G D D K T T G F R A F L A R I M P Q -  
  
TACGTCCGTCATGGTCCCCACCGTGGACCGGGCGTTGCCCCCACGGGCTTCTGGTCGAC  
9541 -----+-----+-----+-----+-----+-----+-----+ 9600  
ATGCAGGCAGTACCAGGGGTGGCACCTGGCCCCGAACGGGGGTGCCCGAAGACCAGCTG  
-1 V D T M T G V T S R A N G G V P K Q D V -  
  
GATCACCGGGGTGGTGGAGGTTCTCGATCGCCTCGGCCTGAGGACGTTCTGTACTTCGAAG  
9601 -----+-----+-----+-----+-----+-----+-----+ 9660  
CTAGTGGCCCCACCACTCCAAGAGCTAGCGGAGCCGGACTCCTGCAAGCATGAAGCCTTC  
-1 I V P T T L N E I A E A Q P R E Y K P L -  
  
CTGGTTGCGGATGTACCAGCTGAAGGTGGAGTTCAGCTGTGCTGGGCCTCCACGGCCAC

```

9661 -----+-----+-----+-----+-----+-----+ 9720
GACCAACGCCTACATGGTTCGACTTCCACCTCAAGTCGACAGCGACCCGGAGGTGCCGGTG
-1      Q N R I Y W S F T S N L Q R Q A E V A V -

CGTGTCTGAAGACGATCGACGACTTGCCCGAACCCGAGACCCCGTGAAGACCGTGATCTG
9721 -----+-----+-----+-----+-----+-----+ 9780
GCACAGCTTCTGCTAGCTGCTGAACGGGCTTGGGCTCTGGGGGCACTTCTGGCACTAGAC
-1      T D F V I S S K G S G S V G T F V T I Q -

GTTGCGGGGAATCGTCAGGGAGACATCTTTGAGGTTGTGGATCCGCGCGCCCGCGATGCG
9781 -----+-----+-----+-----+-----+-----+ 9840
CAACGCCCCCTTAGCAGTCCCTCTGTAGAACTCCAACACCTAGGCGCGCGGGGCGCTACGC
-1      N R P I T L S V D K L N H I R A G A I R -

GATGCCGTCTCCCGGGCCGGATGTTTTTCCCGCGCCGGCGGTGGGGTCCGGTGACGCTCAC
9841 -----+-----+-----+-----+-----+-----+ 9900
CTACGGCAGAGGGCCCCGGCCTACAAAAAGGGCGCGGCCGCCACCCAGCCACTGCGAGTG
-1-<    I G D G P G S T K G A G A T P D T V S M-

AGAGTTTCTCCTGGCTTCCGTACATGATTTACCGTGTCTAGCCGGGCAAACCGGCGGAA
9901 -----+-----+-----+-----+-----+-----+ 9960
TCTCAAAAAGGAGGACCGAAGGCATGTACTAAATGGCACAGTCGGCCCGTTTGGCCGCCTT

CGGTAACCACCTAGCTTGTACTCAGGAGGTGTCCGGGGTCTTCTCCTCCCGTGCTGACTT
9961 -----+-----+-----+-----+-----+-----+ 10020
GCCATTGGTGGATCGAACATGAGTCCTCCACAGGCCCCAGAAGAGGAGGGGCACGACTGAA
0-*      * S T D P T K E E R A S K -

GGGGGCCGGCCCCGCGGACAGGGCCGGCTCCGTGTTCCACCCCGCCAGCCGATCCCCCGG
10021 -----+-----+-----+-----+-----+-----+ 10080
CCCCCGGCCGGGGCGGCTGTCCCGGCCAGGCACAAGGTGGGGCGGTCCGGCTAGGGGGGC
0      P A P G G S L A P E T N W G A L R D G R -

CTCCGTCTCGTCCTCCTCGAGAACGATCCGGCTGCTCGCCCAGCGCAGGATCGGCGGCGC
10081 -----+-----+-----+-----+-----+-----+ 10140
GAGGCAGAGCAGGAGGAGCTCTTGCTAGGCCGACGAGCGGGTCCGCTCCTAGCCGCCGCG
0      E T E D E E L V I R S S A W R L I P P A -

CGTCACCGAGGTGATGAGGGCGACCAGCACGATGATCGTGAAGGTCACGGTGTCCAGTAC
10141 -----+-----+-----+-----+-----+-----+ 10200
GCAGTGGCTCCACTACTCCCGCTGGTCTGCTACTAGCACTTCCAGTGCCACAGGTCATG
0      T V S T I L A V L V I I T F T V T D L V -

GCCGATACGCAGGCCGACCAGGGCGATCACCACCTCGATCATTCCACGCGAGTTCATCCC
10201 -----+-----+-----+-----+-----+-----+ 10260
CGGCTATGCGTCCGGCTGGTCCCGCTAGTGGTGGAGCTAGTAAGGTGCGCTCAAGTAGGG
0      G I R L G V L A I V V E I M G R S N M G -

CGCTCCGAGCGCCAGCCCCTCGTAGCGGCTCATCCCGCCACTACGGGCGGCGACGTACGC
10261 -----+-----+-----+-----+-----+-----+ 10320
GCGAGGCTCGCGGTCCGGGAGCATCGCCGAGTAGGGCGGTGATGCCCGCCGCTGCATGCG
0      A G L A L G E Y R S M G G S R A A V Y A -

ACCGGCGAACTTGCCGAAAGTGGCCACCAACAGCACCCCCGAGGCCCGTGAGCAGCACCGA
10321 -----+-----+-----+-----+-----+-----+ 10380
TGGCCGCTTGAACGGCTTTTACCAGGTGGTTGTGCTGGGGCTCCGGGCACTCGTCTGGCT
0      G A F K G F T A V L L V G L G T L L V S -

CGGCTCCGCGAGTGCGGTCCAGGTCCATGCGAAGCCCCCACTGCCAGGAACACCGGTGC
10381 -----+-----+-----+-----+-----+-----+ 10440
GCCGAGGCGCTCACGCCAGTCCAGGTACGCTTCCGGGTGTGACGGGTCTTGTGGCCACG
0      P E A L A T L D M R L G V S G L F V P A -

GAACACGGCCATGACCAGCGTGCGCAGCGGGGCGAGCCGTACCGGGGCGATGTGCCTCAG
10441 -----+-----+-----+-----+-----+-----+ 10500
CTTGTGCCGGTACTGGTTCGACGCGTCGCCCCGCTCGGCATGGCCCCGCTACACGGAGTC

```

0 F V A M V L T R L P A L R V P A I H R L -  
 CAGGGTTCGACCCGGCCACGAACGCCCCGAACAACGCCTCCATCCCGGCCCGCCGCGGTACG  
 10501 -----+-----+-----+-----+-----+ 10560  
 GTCCACAGCGTGGCCGGTGCTTGCGGGGCTTGTTGCGGAGGTAGGGCCGGCGGCCAGTC  
 0 L T A G A V F A G F L A E M G A A A T L -  
 CGCCCCGTACAGGACGACCACGGCCACGCCGACGGTGACGGCCGATACGGGGACCCGGCT  
 10561 -----+-----+-----+-----+-----+ 10620  
 GCGGGGCATGTCTGCTGGTGCCGGTGCGGCTGCCACTGCCGGCTATGCCCTGGGGCCGA  
 0 A G Y L V V V A V G V T V A S V P V R S -  
 GTCACCCGTACGGGACAGCCGCTGCCGATCGGGCCGCCACCGCACACGCCGCGGCGAC  
 10621 -----+-----+-----+-----+-----+ 10680  
 CAGTGGGCATGCCCTGTGCGCGGACGGCTAGCCCGGCGGGTGGCGTGTGCGGCGCCGCTG  
 0 D G T R S L R R G I P G G V A C A A A V -  
 GAAGACGGTTCGTCCAGGCCATCGTGGTCAGGACCACGGGGCCCCCGGCCGCCCACTCGC  
 10681 -----+-----+-----+-----+-----+ 10740  
 CTTCTGCCAGCAGGTCCGGTAGCACCAGTCTTGGTGCCCGGGGGCGGCGGGGTGAGCG  
 0 F V T T W A M T T L V V P G G A A G S A -  
 CAGCGCCGTACACAGAGCGAGCAGCAGCCAGCCACCGCGTCGTGAAACACCGCTGCCGC  
 10741 -----+-----+-----+-----+-----+ 10800  
 GTCGCGGCAGTGGTCTCGCTCGTCTGCTCGGTGCGGTGGCGCAGCAGCTTGTGGCGACGGCG  
 0 L A T V L A L L L W G V A D D F V A A A -  
 GATGAGCAGCTGGCCGACGTTGCGGTGCGTCAGATTAGGTTCGGCGAGCGTCTTGGCGAT  
 10801 -----+-----+-----+-----+-----+ 10860  
 CTACTCGTCGACCGGCTGCAACGCCACGCAGTCTAAGTCCAGCCGCTCGCAGAACCGCTA  
 0 I L L Q G V N R H T L N L D A L T K A I -  
 CACCGGGAGGGCCGTGACACACATCGCGACCCCGAGGAACAGCGCGAAGACGCCCCGCTC  
 10861 -----+-----+-----+-----+-----+ 10920  
 GTGGCCCTCCCGGCACTGTGTGTAGCGCTGGGGCTCCTTGTGCGGCTTCTGCGGGGCGAG  
 0 V P L A T V C M A V G L F L A F V G R E -  
 TCCGGAGTCCGCGAGCAGCGAGGCGGGCACCAGGTAGCCGGTGGCGATGCCAGCCCCAG  
 10921 -----+-----+-----+-----+-----+ 10980  
 AGGCCTCAGGCGCTCGTCTCGTCCGCCCCGTGGTCCATCGGCCACCGCTACGGGTGCGGGTC  
 0 G S D A L L S A P V L Y G T A I G L G L -  
 AGGAATCAGAAGACCCGCCAGGCTGACCCGGGCGGCCAGACCCCGCGCTTGCAGCAGGAT  
 10981 -----+-----+-----+-----+-----+ 11040  
 TCCTTAGTCTTCTGGGCGGTCCGACTGGGCCCCGCGGTCTGGGGGCGCGAACGCGTCCTA  
 0 P I L L G A L S V R A A L G G R K R L I -  
 CCGGGGGTTCGAAC TGGGCACCTGCGATGGCCACCAGCAGAAGGACGCCGAAC TGGCAGAA  
 11041 -----+-----+-----+-----+-----+ 11100  
 GGCCCCCAGCTTGACCCGTGGACGCTACCGGTGGTCTCTTCTGCGGCTTGACCGTCTT  
 0 R P D F Q A G A I A V L L L V G F Q C F -  
 CGCGTCGAGCAGGTGCGCCTGCGAGATGTCTCGGAAACAGCCTGCCGAAAGTCCCGG  
 11101 -----+-----+-----+-----+-----+ 11160  
 GCGCAGCTCGTCCACGCGGACGCTCTACAGGAGCCCTTTGTGCGACGGCCCTTTCAGGGCC  
 0 A D L L H A Q S I D E P F L R G S L G P -  
 CGAGATCTGCCCCAGCAGGGTCGGCCCCGAGCAGTACCCCCGCGGTGAGCTCCCCACCAG  
 11161 -----+-----+-----+-----+-----+ 11220  
 GCTCTAGACGGGGTCTGTCAGCCGGGCTCGTCATGGGGGCGCCAGTCGAGGGGGTGGTC  
 0 S I Q G L L T P G L L V G A T L E G V L -  
 CGGCGGCAGACCGATCCGGGTCCCCAGCCGTCCAGACCGTAGGCACAGGCGAGCAGGAG  
 11221 -----+-----+-----+-----+-----+ 11280  
 GCCGCCGTCTGGCTAGGCCCAGGGGTCCGCGAGGGTCTGGCATCCGTGTCCGCTCGTCCTC  
 0 P P L G I R T G L R G L G Y A C A L L L -

094545-04500

```

GCCGACCTGGAGCAGGAAGACCGTCAGCGGCTCCCCGCCAGCGGCGACGTGGCTGCGAG
11281 -----+-----+-----+-----+-----+-----+ 11340
CGGCTGGACCTCGTCCTTCTGGCAGTCGCCGAGGGGCGGGTCGCCGCTGCACCGACGCTC
0      G V Q L L F V T L P E G G L P S T A A L -

CACAGCCACGTACAGGACCGCGCACCGGGAACCCAGCCCAGCCCGTCCGTCGACGCGGCCA
11341 -----+-----+-----+-----+-----+-----+ 11400
GTGTCGGTGCAGTCCTGGCGCGTGGCCCTTGGGTGCGGGTCGGGCAGGCAGCTGCGCCGGT
0-<      V A V -
11-*      * S R A G P V W G L G D T S A A -

GACCCCCCTGCCTCACCGGTGCTCGGCCCGCCCTCATCCCCAGAAGAGCCCGTGCCT
11401 -----+-----+-----+-----+-----+-----+ 11460
CTGGGGGGACGGAGTGGCCAGCGAGCCGGGGGCGGAGTAGGGGGTCTTCTCGGGCACGGA
11      L G G Q R V P R E A G A E D G L L A R A -

GCAGTGCGGCGCTCTGCTCCATGAGGCGGCCACCACCTTTCCCGGCACGGCGCCGTGCG
11461 -----+-----+-----+-----+-----+-----+ 11520
CGTCACGCCCGAGACGAGGTACTCCGCCGGGTGGTGGAAAGGGCCGTGCCGCGGCACGC
11      Q L A A S Q E M L R G V V K G P V A G H -

GCCCCGTCGGCGTCGCCCCGAGCGGTGTGCGTCATGCCGGCCATCTCGTCGGACGCCTCGG
11521 -----+-----+-----+-----+-----+-----+ 11580
CGGGCAGCCCGCAGCGGGCGTCGCCACACGCAGTACGGCCGGTAGAGCAGCCTGCGGAGCC
11      P G D A D G A A T H T M G A M E D S A E -

AGAACCGCTGCCTGGCCCCGGGCCGTGTGCGGCAACTCGTCGGAGGAGACCCCGCCGATCA
11581 -----+-----+-----+-----+-----+-----+ 11640
TCTTGGCGACGACCGGGCCCGGCACAGCCGCTTGAGCAGCCTCCTCTGGGGCGGCTAGT
11      S F R Q R A R A T D A F E D S S V G G I -

GTTTCGACGAAGGACTGCAGGTCCGAGTCCGCGGTGTTGGAGATCTTCCGGGCCTGCCAGA
11641 -----+-----+-----+-----+-----+-----+ 11700
CAAGCTGCTTCTGACGTCCAGCCTCAGGCGCCACAACCTCTAGAAGGCCCGACGGTCT
11      L E V F S Q L D S D A T N S I K R A Q W -

AATAGGAGTCCTCCGAATGGTGCATGTCTGATAGAAGCCGACCAGGAACCTGATAGAAGCGGC
11701 -----+-----+-----+-----+-----+-----+ 11760
TTATCCTCAGGAGGCTTACCACGTACAGCATCTTCCGCTGGTCTTGGAGCATCTTCGCCG
11      F Y S D E S H H M D Y F G V L F E Y F R -

CGTACTCCAGCCGGTAGCGGGCCTCGAACTCCTCGAACCGCGCTGGTCTCGTCGACCGACC
11761 -----+-----+-----+-----+-----+-----+ 11820
GCATGAGGTGCGCCATCGCCCGGAGCTTGAGGAGCTTGCGCGACCAGAGCAGCTGGCTGG
11      G Y E L R Y R A E F E E F A S T E D V S -

CGTCCAGGCAGGAGTTGAGCGAGCGCGCTGCCAGCAGTCCGCTGTAGGTGGCGAGGTGCA
11821 -----+-----+-----+-----+-----+-----+ 11880
GCAGGTCCGTCTCAACTCGCTCGCGCGACGGTCGTACGGCGACATCCACCGCTCCACGT
11      G D L C S N L S R A A L L G S Y T A L H -

CCCCGGAGGAGAACACCGGGTCGACGAAGCACGCGGCATCCCCGACCAGGGCCATGCCCCG
11881 -----+-----+-----+-----+-----+-----+ 11940
GGGGCCTCCTCTTGTGGCCAGCTGCTTCTGTCGCCGTAGGGGCTGGTCCCGGTACGGGC
11      V G S S F V P D V F C A A D G V L A M G -

GCGCCCAGAACTTCGTGTTGCTGTACGACCAGTCCTTGCGGACCCGGAGCTCGCCGTAGG
11941 -----+-----+-----+-----+-----+-----+ 12000
CGCGGGTCTTGAAGCACAACGACATGCTGGTCAGGAACGCCTGGGCCTCGAGCGGCATCC
11      P A W F K T N S Y S W D K R V R L E G Y -

GGCCCTCGGTACCCCGGTGGCCTCGGAGAGCTTCTCCGCGATCAGCGGGCAGGCCGCGA
12001 -----+-----+-----+-----+-----+-----+ 12060
CCGGGAGCCAGTGGGCCCACCGGAGCCTCTCGAAGAGGCGCTAGTCGCCCGTCCGGCGCT
11      P G E T V R T A E S L K E A I L P C A A -

```

	TGAACGACTCCATCGCCTTCTCGGGGTGCACCAGGCTCGCCGAGTCCCgggTTCA	
11	12061 -----+-----+-----+-----+-----+-----+-----+ 12120 I F S E M A K E P D G Q V L S A S D R N -  CCACTGCGCCGACACTCGTCAGCTCGGGAGACAGGGGTATGTACCAGAACCACCCGTGCT	
	12121 -----+-----+-----+-----+-----+-----+-----+ 12180 GGTGACGCGGCTGTGAGCAGTCGAGCCCTCTGTCCCCATACATGGTCTTGGTGGGCACGA	
11	V V A G V S T L E P S L P I Y W F W G H -  CGAAGGTGCAGGTGAAGATGTTCCCGGAGTTTCGGCTTCGGAAGCCGCTTGCCGCCGTTGA	
	12181 -----+-----+-----+-----+-----+-----+-----+ 12240 GCTTCCACGTCCACTTCTACAAGGGCCTCAAGCCGAAGCCTTCGGCGAACGGCGGCAACT	
11	E F T C T F I N G S N P K P L R K G G N -  AGTAGCCGAACAGGGCCAGGTTGCGGAAGAAGGGCGAGTACTCGCGCTTGCGCGCCGACT	
	12241 -----+-----+-----+-----+-----+-----+-----+ 12300 TCATCGGCTTGTCCCGGTCCAACGCCTTCTTCCCGCTCATGAGCGCGAACC CGGGCTGA	
11	F Y G F L A L N R F F P S Y E R K A G S -  TCTTGTACAGCCCACCGGTGTTGCGGAGGCGTCCACGACGAAACGGGAGCCACCTCGT	
	12301 -----+-----+-----+-----+-----+-----+-----+ 12360 AGAACATGTCGGGTGGCCACAACGGCCTCCGCAGGTGCTGCTTTGCCCTCGGGTGGAGCA	
11	K K Y L G G T N G S A D V V F R S G V E -  GCTCGCGCCCCCTCGGAGTCCCGGTAGCGCACGCCCCGCACCCGGCCGTCCTCGGCCTTGA	
	12361 -----+-----+-----+-----+-----+-----+-----+ 12420 CGAGCGCGGGGAGCCTCAGGGCCATCGCGTGCGGGGCGTGGGCGCGCAGGAGCCGGAAct	
11	H E R G E S D R Y R V G R V R G D E A K -  GCACGTCGAGGACATCGCTGTTCTCCCGCACCTCGACACCGTGCCTGCGAGCGTTGTCTGA	
	12421 -----+-----+-----+-----+-----+-----+-----+ 12480 CGTGCAGCTCCTGTAGCGACAAGAGGGCGTGGAGCTGTGGCACGGACGCTCGCAACAGCT	
11	L V D L V D S N E R V E V G H R R A N D -  GCAGGATCTGGTCTGAACCTCATGCGCTCGACCTGGTACGCGTACCCCGTCGCCCCCGGCA	
	12481 -----+-----+-----+-----+-----+-----+-----+ 12540 CGTCC TAGACCAGCTTGAAGTACGCGAGCTGGACCATGCGCATGGGGCAGCGGGGGCCGT	
11	L L I Q D F K M R E V Q Y A Y G T A G P -  TCCGCGCGGAGACGGCGAAGTCTGAACGTCCACGGTTTCGGGGTTGGCACCCCACTTGAACG	
	12541 -----+-----+-----+-----+-----+-----+-----+ 12600 AGGCCGCGCTCTGCCGCTTCAGCTTGCAGGTGCCAAGCCCCAACC GTGGGGTGAAC TTGC	
11	M R R S V A F D F T W P E P N A G W K F -  TCCCGCCGTGCTTGATCGTGAAGGCTGCCCTTCTTCAGCTCGTTCGAGACACCGAGGAGGT	
	12601 -----+-----+-----+-----+-----+-----+-----+ 12660 AGGGCGGCACGAACTAGCACTTCCGACGGAAGAAGTCGAGCAGCCTCTGTGGCTCCTCCA	
11	T G G H K I T F A A K K L E D S V G L L -  GTGCGATGCCGTGGACGGTGGAGGGGAGGAGCGACTCACCGATCTGGTAGCGCGGGAAGG	
	12661 -----+-----+-----+-----+-----+-----+-----+ 12720 CACGCTACGGCACCTGCCACCTCCCTCCTCGCTGAGTGGCTAGACCATCGCGCCCTTCC	
11	H A I G H V T S P L L S E G I Q Y R P F -  TCTCCTTCTCCAGCTGGAGTACGCGATGGCCCCGCTTGCGGACCAGCGTGGAGACGGTCG	
	12721 -----+-----+-----+-----+-----+-----+-----+ 12780 AGAGGAAGAGGTCGACCTCATGCGCTACCGGGCGAACGCCTGGTCGCACCTCTGCCAGC	
11	T E K E L Q L V R H G R K R V L T S V T -  AGCCCCGCCGACCTCCGCCGACCACGATGACGTCGTACTGCGCTGACACGTCCACGGACT	
	12781 -----+-----+-----+-----+-----+-----+-----+ 12840 TCGGGCGGCCTGGAGGCGGCTGGTGCTACTGCAGCATGACGCGACTGTGCAGGTGCCTGA	
11-<	S G A P G G G V V I V D Y Q A S V D M -  CTCCTTCTCGCACATCGGGCGTCTCATATTCCAGGAATCCTCTGGCCCCGCCAGGTGCT	





```

GTCGAGGACGTCGTGGAGTGCAACCGGCGGATGCTCGGCCGTCTGGCGCTCCAGGTGTCTG
13681 -----+-----+-----+-----+-----+-----+ 13740
CAGCTCCTGCAGCACCTCACGTTGGCCGCCTACGAGCCGGCAGACCGCGAGGTCCACAGC
10 V E D V V E C N R R M L G R L A L Q V S -

GGCGAGGTGGACCCGGAGAGCGAACTGGTGGGTGCGGTGGTTCGTGAGGAGGGCGCCCGG
13741 -----+-----+-----+-----+-----+-----+ 13800
CCGCTCCACCTGGGCCTCTCGCTTGACCACCCACGCCACCAGCAGCTCCTCCCGCGGGCC
10 G E V D P E S E L V G A V V V E E G A R -

GTGACGCGTTTCGCGGGTTCGTGGGACCAGCGGTGATCGGCGCGGGCACGGTTCGTGAGGAC
13801 -----+-----+-----+-----+-----+-----+ 13860
CACTGCGCAAGCGCCAGCACCTTGGTCCGCACTAGCCGCGCCCGTGCCAGCAGCTCCTG
10 V T R S R V V G P A V I G A G T V V E D -

AGCCAGATCGGACCGTACGCCTCCATCGGCCGGCGCTGCACCGTGCGGGCGTCCCGGCTC
13861 -----+-----+-----+-----+-----+-----+ 13920
TCGGTCTAGCCTGGCATGCGGAGGTAGCCGGCCGCGACGTGGCACGCCCGCAGGGCCGAG
10 S Q I G P Y A S I G R R C T V R A S R L -

TCCGACTCCATCGTCCTTGACGACGCCTCGATCCTCGCGGTGAGCGGACTGCACGGCTCG
13921 -----+-----+-----+-----+-----+-----+ 13980
AGGCTGAGGTAGCAGGAAGTCTGCGGAGCTAGGAGCGCCACTCGCCTGACGTGCCGAGC
10 S D S I V L D D A S I L A V S G L H G S -

CTGATCGGAAGGGGCGCGGGATCGCGCCCGGGGCCCGGGGCGAGGCCCGGCACCGGCTG
13981 -----+-----+-----+-----+-----+-----+ 14040
GACTAGCCTTCCCGCGCGCCTAGCGCGGGCCCGGGGCCCGCTCCGGGCGGTGGCCGAC
10 L I G R G A R I A P G A R G E A R H R L -

GTCGTGCGGCACACGTGCAGATCGAGATCGCGGCCTGACGCACCCACCGGAGCACCGGG
14041 -----+-----+-----+-----+-----+-----+ 14100
CAGCAGCCGCTGGTGCACGTCTAGCTCTAGCGCCGACTGCGTGGGTGGCCTCGTGGCCC
10-* V V G D H V Q I E I A A * -

GGGAGGCTCGGCAGGGGCGTCAGGCCGTAAGAAGGGCTGCCGGGGCGGGACGGACCCGCC
14101 -----+-----+-----+-----+-----+-----+ 14160
CCCTCCGAGCCGTCCCGCAGTCCGGCATTCTTCCCGACGGCCCCGCCCTGCCTGGGCGG

CCGGCAGCCACAGGTCCCCGGTCCGCGGATATGGGGGACTCGAGGTTTCGATCAGCCGAA
14161 -----+-----+-----+-----+-----+-----+ 14220
GGCCGTCGGGTGTCCAGGGGCCAGGCGCCTATACCCCTGAGCTCCAAGCTAGTCGGCTT
9-* * G F -

GGTCAGAGCCACGTGGCCGAGGTTCGAGCCCGGAGTTGCCGGCGCCGAGGTTACAGGCGGC
14221 -----+-----+-----+-----+-----+-----+ 14280
CCAGTCTCGGTGCACCGGCTCCAGCTCGGGCCTCAACGGCCGCGGCTCCAATGTCCGCCG
9 T L A V H G L D L G S N G A G L N C A A -

CGTGGCGCAGTCGACGCTGCCGACCGGCGTGCCTTCGGGCGTGGAGCCCGTGTACGACTT
14281 -----+-----+-----+-----+-----+-----+ 14340
GCACCGCGTCAGCTGCGACGGCTGGCCGCACGGAAGCCCGCACCTCGGGCACATGCTGAA
9 T A C D V S G V P T G E P T S G T Y S K -

GCGCACGACGAAGCTGAACGACGCCGCTCCGGACGCGTCCGTGGTGAAGGACGTTCGCGGT
14341 -----+-----+-----+-----+-----+-----+ 14400
CGCGTGCTGCTTCGACTTGCTGCGGCGAGGCCTGCGCAGGCACCACTTCCTGCAGCGCCA
9 R V V F S F S A A G S A D T T F S T A T -

CGCCGGGTTGCACGCGTCTTGCCACCGACCGGAGCGCACTGGGCGATGTAGTAGGTCTC
14401 -----+-----+-----+-----+-----+-----+ 14460
GCGGCCCAACGTGCGCAGGACCGGTGGCTGGCCTCGCGTGACCCGCTACATCATCCAGAG
9 A P N C A D Q G G V P A C Q A I Y Y T E -

GCCGGCGGCGGCACCGCTGACCGACACCGACACGCTCTGTCCGTCACTCAGACCCGAGGC

```

```

14461 -----+-----+-----+-----+-----+-----+ 14520
CGGCCGCGCCGCTGGCGACTGGCTGTGGCTGTGCGAGACAGGCAGTGAGTCTGGGCTCCG
9      G A A A G S V S V S V S Q G D S L G S A -

GGGACTGACGGAGAAGGCGGGCGCGGCGAAGGCGACGGACTGTGCGGCGGCGGCCAGGCC
14521 -----+-----+-----+-----+-----+-----+ 14580
CCCTGACTGCCTCTTCCGCCCCGCGCCGCTTCCGCTGCCTGACACGCCGCCGCGGTCCGG
9      P S V S F A P A A F A V S Q A A A A L G -

GATGGATGCGACGGCCACGACGCCGAACCTGGAAGCACGGCGGGACATGTGACGTAACGA
14581 -----+-----+-----+-----+-----+-----+ 14640
CTACCTACGCTGCCGGTGCTGCGGCTTGGACCTTCGTGCCGCCCTGTACACTGCATTGCT
9      I S A V A V V G F R S A R R S M H R L S -

CATGCGTAGGCTCCGATTTCGAGGAGGGGGTTGATCACTCCATGAAAGGATCACCTCGCCG
14641 -----+-----+-----+-----+-----+-----+ 14700
GTACGCATCCGAGGCTAAGCTCCTCCCCCACTAGTGAGGTACTTTCCTAGTGAGCGGC
9-<      M -
8-*                                     * R A -

GACGGCCCGCCTGCATCTCCCTCTGTGCTCTCGTGGATTTCGGGCACGGCACTCCCGTCCA
14701 -----+-----+-----+-----+-----+-----+ 14760
CTGCCGGCGGACGTAGAGGGAGACACGAGAGCACCTAAAGGCCGTGCCGTGAGGGCAGCT
8      P R G G A D G E T S E H I E P V A S G D -

CGGCCGCGCCGAGAATGCGGCAGACCCCCGCACCTCCTCCGGCCCCACCGCCGTACCGG
14761 -----+-----+-----+-----+-----+-----+ 14820
GCCGGCGGGCGTCTTACGCCGTCTGGGGGGCGTGGAGGAGGCCGGGGTGGCGGCATGGCC
8      V A A R L I R C V G R V E E P G V A T G -

TGGGCAGCGACAGCACCCGCTCGGTGAGCGCCTCCACCTTCGGGAGCGGATCGGGCGCGT
14821 -----+-----+-----+-----+-----+-----+ 14880
ACCCGTCGCTGTCTGGGCGAGCCACTCGCGGAGGTGGAAGCCCTCGCCTAGCCCCGCGCA
8      T P L S L V R E T L A E V K P L P D P A -

GGCGCGCGAGGTTCGGACCGGTAGGGCTCGCAGCTGTGGCAGCCGGGGCTGAAGTAGGCGC
14881 -----+-----+-----+-----+-----+-----+ 14940
CCGCGCGCTCCAGCCTGGCCATCCCAGCGCTCGACACCGTCGGCCCCGACTTCATCCGCG
8      H R A L D S R Y P E C S H C G P S F Y A -

GGGCCAGGACGTTGTGCCGTTGGAGCACCGCCTGGAGTTCGTGCGGTCGAGCCCGGCGC
14941 -----+-----+-----+-----+-----+-----+ 15000
CCCGGTCCTGCAACACGGCAACCTCGTGGCGGACCTCAAGCAGCGCCACGTCGGGCCGCG
8      R A L V N H R Q L V A Q L E D R H L G A -

GGACGGCGTCCACCTCGATGACGACGTAAGTGGCAGTTCGACAGCTCGTTCGGATCCTGCG
15001 -----+-----+-----+-----+-----+-----+ 15060
CCTGCCGCGAGGTGGAGCTACTGCTGCATGACCGTCAAGCTGTGAGCAAGCCTAGGACGC
8      R V A D V E I V V Y Q C N S L E N P D Q -

GGCGGACCCGACGCGCGGCAGTCCGTCGAGGTACTGCTCGTACAGACGGTAGTTGCGCC
15061 -----+-----+-----+-----+-----+-----+ 15120
CCGCTGGGCGTGGCGCCGTCAGGCAGCTCCATGACGAGCATGTCTGCCATCAACGCGG
8      P R V R V G P L G D L Y Q E Y L R Y N R -

GGTTGATCGCGGTGAAGTGATCGGCGGACTCCAGGGAGGTGAGGCCCATGGCCGCGCTGA
15121 -----+-----+-----+-----+-----+-----+ 15180
CCAAGTAGCGCCACTTCACTAGCCGCTGAGGTCCCTCCACTCCGGGTACCGGCGCGACT
8      R N I A T F H D A S E L S T L G M A A S -

TCTCGTGCATCCGCGCGACCGTTCCGCTCCCGGTGATCTCATGCGCGGCGTTGAGCCCCT
15181 -----+-----+-----+-----+-----+-----+ 15240
AGAGCACGTAGGCGCGCTGGCAAGGCGAGGGCCACTAGAGTACGCGCCGCAACTCGGGGA
8      I E H M R A V T G S G T I E H A A N L G -

GGTGGCGCATGGCCCGGAGCCGGTCCGCCAGGGCGTCTCGTGGTGACGATCGCCCCGC

```





```

6      GCGAGGAACGCCTGCTGGAGCAGCCACATCAAGGTCCGCTCCTCCTCCATCATCAGGCCG
      R E K R V V E D T Y N W A L L L Y Y D P -

16801  TCGGTGGCAGCGACCTCCTCCGGAGGAAGGACCGGGATGCGGTTCCCCGGCAGCAGTTTG
      +-----+-----+-----+-----+-----+
6      AGCCACCGTCTGCTGGAGGAGGCCTCCTTCCTGGCCCTACGCCAAGGGGCCGTCGTCAAAC
      E T A A V E E P P L V P I R N G P L L K -

16861  CCGTGCTTGAGGCTGGTGGTGTGCGCGCAGACGGTGATGTCTTGATCCGTGACACCGCAG
      +-----+-----+-----+-----+-----+
6      GGCACGAACCTCCGACCACCACAGCGGCGTCTGCCACTACAGGACTAGGCAGTCTGGCGTC
      G H K L S T T D G C V T I D Q D T L G C -

16921  GCCATCAGCAACTGGGTCCCCTTGAGCGGTGCTCCGTAGCCGGCCACGCGGTGGCCGTCC
      +-----+-----+-----+-----+-----+
6      CGGTAGTCGTTGACCCAGGGGAACCTGCCACGAGGCATCGGCCGCTGCGCCACCGGCAGG
      A M L L Q T G K S P A G Y G A V R H G D -

16981  GCGGCCAGACCGCGAACGAGCGTACGGATCGCTTCGGTCACGCGCGTACCCGCTCGGCG
      +-----+-----+-----+-----+-----+
6      CGCCGGTCTGGCGCTTGCTCGCATGCCTAGCGAAGCCAGTGCAGCGAGTGGGCGAGCCGC
      A A L G R V L T R I A E T V R T V R E A -

17041  AACCCCCGGTAGGGGGCATCCGTGACAGTCCGCGCTCCTCCTCAGGCCGAGCAGCGCC
      +-----+-----+-----+-----+-----+
6      TTGCGGGCCATCCCCCGTAGGCAGTCGTGAGCGCGAGGAGGAGGTCCGGCTCGTCGCGG
      F A R Y P A D T L L G R E E E L G L L A -

17101  GCGACCGAGGGCTCCGGGACCCGTGCGGCCGACTCGCGCGCGGCGACGACCGCGATCGAA
      +-----+-----+-----+-----+-----+
6      CGCTGGCTCCCCAGGCCCCTGGGCACGCCGGCTGAGCGCGCGCCGCTGCTGGCGCTAGCTT
      A V S P E P V R A A S E R A A V V A I S -

17161  CCGCCGTGCACGGCGACCCGCTCCACGTGATGATCCGCAGGCCGTGCGCGCCGAAGAGG
      +-----+-----+-----+-----+-----+
6      GGCGGCACGTGCCGCTGGGCGAGGTGCAGCTACTAGGCGTCCGGCACGCGCGCTTCTCC
      G G H V A V R E V D I I R L G H A G F L -

17221  TGGCGCAGTGTGTGACGGGAGAAGTACGACAGGTGCTCGTGGTAGATCGTGTGCAACTGG
      +-----+-----+-----+-----+-----+
6      ACCGCGTCACACAGTCCCTCTTCATGCTGTCCACGAGCACCATCTAGCACAGCTTGACC
      H R L T H L S F Y S L H E H Y I T D F Q -

17281  TTCTCGTCGAGCAGGTTTCAGCAGGTACGGCACCTCGATGACCAGGACGCCGTGCTGCTCG
      +-----+-----+-----+-----+-----+
6      AAGAGCAGCTCGTCCAAGTCGTCCATGCCGTGGAGCTACTGGTCTGCGGCAGCAGCAGC
      N E D L L N L L Y P V E I V L V G D D D -

17341  AGCACTGCGTCGACGCCGTCCAGGATGCGGTGCACGTGCTGATGTGCGCGAAGCACTGG
      +-----+-----+-----+-----+-----+
6      TCGTGACGCAGCTGCGGCAGGTCTACGCCACGTGCAGCAGCTACACGCGCTTCGTGACC
      L V A D V G D L I R H V D D I H A F C Q -

17401  CGGCCGATGACGGCCTTGCCCTGCTCAAGGGCGATGCGGCCCGCGGGCTCCGGG
      +-----+-----+-----+-----+-----+
6      GCCGGCTACTGCCGAACCGGGACGGGACGAGTTCCTGCTACGCCGGGCGCCCGAGGCC
      R G I V A K A R G Q E L A I R G A P E P -

17461  CCGAAGAAGTCCGGGTCCGTGGGGATCCCCCGGGCGTTGGCGATCTCGGCGAGGTTGGCC
      +-----+-----+-----+-----+-----+
6      GGCTTCTTCAGGCCCAGGCACCCCTAGGGGGCCCGCAACCGCTAGAGCCGCTCCAACCGG
      G F F D P D T P I G R A N A I E A L N A -

17521  GCCGGGTGACCCCGGCCACCCGCATGCCCGCCCGCCGAACATCGCGAGCTGGGTGCCG
      +-----+-----+-----+-----+-----+
6      CGGCCAGCTGGGGCCGGTGGGCGTACGGGCGGCGGCCTTGCTAGCGCTCGACCCACGGC

```

6        A P D V G A V R M G A A R F M A L Q T G -  
 ACGTTGCTGCCCAGCTCCACGACCAGGTCGCCGAGGCGAGGCTTGCCCGGCGGGTCGCC  
 17581 -----+-----+-----+-----+-----+ 17640  
 TGCAACGACGGGTCGAGGTGCTGGTCCAGCGGCTCCGCTCCGAACGGGCGCCAGCGG  
 6        V N S G L E V V L D G S A L S A R R T A -  
 AGCCCGACGATGTGCGCCATGTGCTCGCGGATCTGGTCGGAGTCGGAGGAGACGTAGACG  
 17641 -----+-----+-----+-----+-----+ 17700  
 TCGGGCTGCTACACGCGGTACACGAGCGCCTAGACCAGCCTCAGCCTCCTCTGCATCTGC  
 6        L G V I H A M H E R I Q D S D S S V Y V -  
 TAGTGCTTGAACAGTGTCCCGGGTCGACGACATGGCGAAGCGTCATCAGCCGGCACGAC  
 17701 -----+-----+-----+-----+-----+ 17760  
 ATCACGAACCTTGTACAGGGCCCCAGCTGCTGTACCGCTTCGAGTAGTCGGCCGTGCTG  
 6        Y H K F L T G P D V V H R L T M L R C S -  
 CGGCACACGATGACGTCGAGCGGGAAGACGTCCTGCGCCTCATCGGCGTCGGCCGGATCG  
 17761 -----+-----+-----+-----+-----+ 17820  
 GCCGTGTGCTACTGCAGCTCGCCCTTCTGCAGGACGCGGAGTAGCCGAGCCGGCCTAGC  
 6        R C V I V D L P F V D Q A E D A D A P D -  
 ACGAACCCGTTGGCCAGCGGCAGCGAGCCGAAGGAGATCACCTCGGTCCAGTCGTCCGCA  
 17821 -----+-----+-----+-----+-----+ 17880  
 TGCTTGGGCAACCGGTCGCCGTCGCTCGGCTTCTCTAGTGGAGCCAGGTCAGCAGGCGT  
 6        V F G N A L P L S G F S I V E T W D D A -  
 CCGCATACACGGCACGTCTCGTCCCGCCTGCATTTCTCCAGCATGAAGTCTCCTGACGGC  
 17881 -----+-----+-----+-----+-----+ 17940  
 GGCGTATGTGCCGTGCAGAGCAGGGCGGACGTAAAGAGGTCGTACTTCAGAGGACTGCCG  
 6-<        G C V R C T E D R R C K E L M -  
 GAATGCCGACGCATCGGGCCCCGTCGGTCCGGGGACGGTCAATCTAGGGTTCCGGCCGACG  
 17941 -----+-----+-----+-----+-----+ 18000  
 CTTACGGCTGCGTAGCCCGGGCAGCCAGGCCCCCTGCCAGTTAGATCCCAAGGCCGGCTGC  
 GCGCTCCACTTCGTATGTGCCCTACTGGTTTCAGCGGAGCGGACGGGTGAACGCCCCGTAC  
 18001 -----+-----+-----+-----+-----+ 18060  
 CCGCGAGGTGAAGCATACAGGGATGACCAAGTCGCCTCGCCTGCCCACTTGGGGCATG  
 17-\*        \* R L P R T F A R V -  
 GTCCTCGATGAGGAGCTGCGGCTGCTCCATGGCCGCGAAGTGCCCGCCGCGGTGGAATC  
 18061 -----+-----+-----+-----+-----+ 18120  
 CAGGAGCTACTCCTCGACGCCGACGAGGTACCGGCGCTTCACGGGCGGCGCCAGCTTGAG  
 17        D E I L L Q P Q E M A A F H G G R D F E -  
 GGTCCACCGCGTCAGGGTCGGCAGGATGCCCTCGGCGAACGACCGGATCGGCCGGGTGGC  
 18121 -----+-----+-----+-----+-----+ 18180  
 CCAGGTGGCGCAGTCCCAGCCGTCCTACGGGAGCCGCTTGCTGGCCTAGCCGGCCCCACCG  
 17        T W R T L T P L I G E A F S R I P R T A -  
 GTCGTCCGGGAACACCGCGACGCCGACGGGGGCGTCAGCGGCCAGGGCCCCGCCCCAGGT  
 18181 -----+-----+-----+-----+-----+ 18240  
 CAGCAGGCCCTTGTGGCGCTGCGGCTGCCCCCGGCAGTCGCCGTCGCCGGCGGGGTCCA  
 17        D D P F V A V G V P A T L P W P G G W T -  
 GCGGGCGAAGTCCGCCATGCCGCGAGCCGACTCGTAGTACAACCTGAGCGCTGGAACCGGC  
 18241 -----+-----+-----+-----+-----+ 18300  
 CGCCCGCTTCAGGCGGTACGGCGCTCGGCTGAGCATCATGTTGACTCGCGACCTTGGCCG  
 17        R A F D A M G R A S E Y Y L Q A S S G A -  
 CGTCGCGGTGAGCCAGTAGATCATCACGTGGGTGAGCAGCCGGTCCCGGGAGATGGCCTC  
 18301 -----+-----+-----+-----+-----+ 18360  
 GCAGCGCCAGTCGGTCATCTAGTAGTGCACCCACTCGTCGGCCAGGGCCCTCTACCGGAG  
 17        T A T L W Y I M V H T L L R D R S I A E -

18361  
17  
18420  
18421  
17  
18480  
18481  
17  
18540  
18541  
17  
18600  
18601  
17  
18660  
18661  
17  
18720  
18721  
17  
18780  
18781  
17  
18840  
18841  
17  
18900  
18901  
17  
18960  
18961  
17  
19020  
19021  
17  
19080  
19081  
17  
19140

CTCCACGTTCTTGGCCGCCGCTCCACTCCTGGAACCTTGTCGAGAATCCAGGCGAGCTGGCC  
GAGGTGCAAGAACGGCGGCGAGGTGAGGACCTTGAACAGCTCTTAGGTCCGCTCGACCGG  
E V N K G G S W E Q F K D L I W A L Q G -  
GACCGGGGAGTTCGGTGAGGCCGTAGGCCAGGGTCTGCGGGCGGGTGGCCTGGATGCGCTG  
CTGGCCCCCTCAGCCACTCCGGCATCCGGTCCCAGACGCCCGCCACCGGACCTACGCGAC  
V P S D T L G Y A L T Q P R T A Q I R Q -  
CCAGCCGATGCCGGTGTTCGGCGAACTCCCCGCTGTGCGCCAGCTTGCCCGAGGTTCGCTCTC  
GGTTCGGCTACGGCCACAGCCGCTTGAGGGGCGACACGCGGTTCGAACGGGTCCAGCGAGAG  
W G I G T D A F E G S H A L K G L D S E -  
GTCCAGGCGCCCCGATGGCCTCCGGGGCGTCTTGGGGCGGGAAGGTCACCAGCATGTTTCAG  
CAGGTCCGCGGGGCTACCGGAGGCCCGCGAGGACCCCGCCCTTCCAGTGGTTCGTACAAGTC  
D L R G I A E P A D Q P P F T V L M N L -  
GTGGACGCCGCCACGTGCTCGGGGTCGGGCCAGCCCCAGCTCCAGCGAGACGACCTTTCC  
CACCTGCGGCCGGTGCACGAGCCCCAGCCGGTTCGGGGTCGAGGTTCGCTCTGCTGGAAAGG  
H V G A V H E P D A L G L E L S V V K G -  
CCAGTCGCCGCCCTGGGCGACGTAACGCTTCGTAGCCGAGGCGGTTTCATCAGCTCCGCCCCA  
GGTCAGCGGCGGGACCCGCTGCATTGCGAGCATCGGCTCCGCCAAGTAGTCGAGGCGGGT  
W D G G Q A V Y R E Y G L R N M L E A W -  
GGCGCGTTCGATCCGCCGCACGTCCCAGCCCGGCTCGGCAGTCGGGCGGAGAAGCCGTA  
CCGCGCACGCTAGGCGCGTGCAGGGTCGGGCCGAGCCGTCAGCCCGGCCCTCTTCGGCAT  
A R A I R R V D W G P E A T P G S F G Y -  
GCCCCGCATGGAGGGGACGACGACGTGGAAGGCGTCCGCCGGGTTCGCCGCCGTGCGCGCG  
CGGGCCGTACCTCCCCTGCTGCTGCACCTTCGCGAGGCGGCCAGCGGCGGCACGCGCGC  
G P M S P V V V H F A D A P D G G H A R -  
CGGGTCGCTCAGCGGCCCGATGACGTCGAGGAACCTCGGCGACCGAGCCCGGCCAGCCGTG  
GCCCAGCGAGTCGCCGGGCTACTGCAGCTCCTTGAGCCGCTGGCTCGGGCCGGTTCGGCAC  
P D S L P G I V D L F E A V S G P W G H -  
GGTGAGGATCAGCGGGATCGCGTCCGGCTCGGGCGAACGCACGTGAAGGAAGTGCACGTC  
CCACTCCTAGTCGCCCTAGCGCAGGCCGAGCCCGCTTGCCTGCACTTCCTTCACGTGCAG  
T L I L P I A D P E P S R V H L F H V D -  
GGCGCCGTCGATCGTGGTGACGAACTGGGGGAACGCGTTTCAGCTCGGCCTCCGCGGCACG  
CCGCGGCAGCTAGCACCCTGCTTGACCCCTTTCGCAAGTCGAGCCGGAGGCGCCGTGC  
A G D I T T V F Q P F A N L E A E A A R -  
CCAGTCGTAGCCGTGGCGCCAGTGGTTCGGTTCGCTCCTTGAGGTAGGACAGCGGCACCTCC  
GGTCAGCATCGGCACCGCGGTTCACCAGCCACTCGAGGAACCTCCATCCTGTGCGCCGTGAGG  
W D Y G H R W H D T L E K L Y S L P V G -  
GCGGTCCCATCCGGATCCGGGTATCTCGGACGGCCACCGGGTTCGCGTCGATCCGCCGGGT  
CGCCAGGGTAGGCCCTAGGCCCATAGAGCCTGCCGGTGGGCCAGCGCAGCTAGGCGGCCCA  
R D W G S G P I E S P W R T A D I R R T -  
TAAGGTCGTTCGAATGTCGGAAGTGGGTTCGATCTCGATACGGAAGGGACGCACAGTGAATCC



19141 -----+-----+-----+-----+-----+ 19200  
ATTCCAGCAGCTTACAGCCTGACCCAGCTAGAGCTATGCCTTCCCTGCGTGTCACTTAGG  
17-< L T T S H R V P D I E I R F P R M -  
ACCCCTCGTGATTGTGGGAGCGGGGCGGCGGAGGCGGCGCCCGATGTGATCCGGGGAC  
19201 -----+-----+-----+-----+-----+ 19260  
TGGGAGACTAACACCCTCGCCCCCGCGCTCCGCGGCGGGGCTACACTAGGCCCTG  
CGTGTCTCAGGCCGGTTCGGCCCGCGCGGCGCGCCTTCCCGTGCAGGAGAAGGACCGCAC  
19261 -----+-----+-----+-----+-----+ 19320  
GCACAGAGTCCGGCCAAGCCGCGCGCGCGGCGGAAGGGCACGCCTCTTCTGGCGTG  
16-\* \* A P E A P A A A G E R A S F S R V -  
GGAGGACAGGAAGTTGCGGATCATCGGCATGCCGTGTTTCGGTCCGGAAGCTCTCCGGATG  
19321 -----+-----+-----+-----+-----+ 19380  
CCTCCTGTCTTCAACGCCTAGTAGCCGTACGGCACAAGCCAGGCCTTCGAGAGGCCTAC  
16 S S L F N R I M P M G H E T R F S E P H -  
GAACTGGACGGACTCCACCGGCAGCGAACGGTGGCGCAGGCCCATCACGTACCCGTCGTC  
19381 -----+-----+-----+-----+-----+ 19440  
CTTGACCTGCCTGAGGTGGCCGTGCTTGCCACCGCGTCCGGGTAGTGCATGGGCAGCAG  
16 F Q V S E V P L S R H R L G M V Y G D D -  
CGTGGAGCGCCCGGTGACCTCGAGGGACGGCGGGACCGTGCCCTCCGGCACGATCAGTGA  
19441 -----+-----+-----+-----+-----+ 19500  
GCACCTCGCGGGCCACTGGAGCTCCCTGCCGCCCTGGCACGGGAGGCCGTGCTAGTCACT  
16 T S R G T V E L S P P V T G E P V I L S -  
GTGGTAGCGGGTCGCGAAGAACCCCGCGGGCAGCCCGGTGAACACTCCGCGCCCGTCGTG  
19501 -----+-----+-----+-----+-----+ 19560  
CACCATCGCCCAGCGCTTCTTGGGGCGCCCGTGGGCCACTTGTGAGGCGCGGGCAGCAC  
16 H Y R T A F F G A P L G T F V G R G D H -  
CGTGATCCGGCTCGTCTTCCCGTGCATGAGATGCCGGGCGGGGACGGTGGCGGCGCCGTA  
19561 -----+-----+-----+-----+-----+ 19620  
GCACTAGGCCGAGCAGAAGGGCACGTACTCTACGGCCCGCCCTGCCACCGCCCGCGCAT  
16 T I R S T K G H M L H R A P V T A A G Y -  
GGCGCGGGCGACGGCCTGATGCCCCAGACAGACCCCGAGCAGCGGGACCCGGCCGGCGAA  
19621 -----+-----+-----+-----+-----+ 19680  
CCGCGCCCGCTGCCGACTACGGGGTCTGTCTGGGGCTCGTCCGCCCTGGGCGCGCCGCTT  
16 A R A V A Q H G L C V G L L P V R G A F -  
GGCCTGGACGATCTCGACGTGCCCCGAGGTGTGCGGGTGGCCGGGGCCCGGCCCCAGCAG  
19681 -----+-----+-----+-----+-----+ 19740  
CCGGACCTGCTAGAGCTGCACGGGCCTCCACAGCCCCACCGCCCCGGGCGGGTCTGTC  
16 A Q V I E V H G S T D P H G P G P G L L -  
GACCGCGTCCGGCCGCATCAGCCCCATCTCGTCCGGGGTCATGAGATGCGACCGCACCAT  
19741 -----+-----+-----+-----+-----+ 19800  
CTGGCGCAGGCCGGCGTAGTCGGGGTAGAGCAGGCCCCAGTACTCTACGCTGGCGTGGTA  
16 V A D P R M L G M E D P T M L H S R V M -  
GACGGGCTCCGCGCCGGCGGACATCAGATACTGGCGCAGGATGTGACGAAGCTGTGAA  
19801 -----+-----+-----+-----+-----+ 19860  
CTGCCCCAGGCGCGGCGCCTGTAGTCTATGACCGGCTCCTACAGCTGCTTCGACAGCTT  
16 V P E A G A S M L Y Q R L I D V F S D F -  
CGCGTCGACCACCAGGACCCGCGGGGCTCGGTGCCTGCGCCGATCCGTGCGGAGACCA  
19861 -----+-----+-----+-----+-----+ 19920  
GCGCAGCTGGTGGTCTTGGGCGCCCCGAGCCACGGACGCGGCCCTAGGCAGCCCTCTGGT  
16 A D V V L V R P A E T G A G S G D P S W -  
CAAGCTCACAGCAACTCCTCTCCGGTGACCGCCAGTGAGTGGCGCTCATCTTGGCCAGC  
19921 -----+-----+-----+-----+-----+ 19980  
GTTTCGAGTGTGTTGAGGAGAGGCCACTGGCGGGTCACTCACCGCGAGTAGAACCGGTG

```

16-<      L  S  M  -
15-*      *  L  L  E  E  G  T  V  A  W  H  T  A  S  M  K  A  L  -

GTCTCGGTCCACTCCGCCCCCGGTTCCGAATCGGCGACGATTCCGGCCGAGGCCCGGGTG
19981 -----+-----+-----+-----+-----+ 20040
CAGAGCCAGGTGAGGCGGGGGCCAAGCCTTAGCCGCTGCTAAGGCCGGCTCCGGGCCAC
15      T  E  T  W  E  A  G  P  E  S  D  A  V  I  G  A  S  A  R  T  -

CGGTAGACGCCCTCGTGGTGGAAAAGGGTCCGGATGCACAGCGCGAGGTTGGTGTACCCG
20041 -----+-----+-----+-----+-----+ 20100
GCCATCTGCGGGAGCACCACCTTTTCCAGGCCTACGTGTGCGGCTCCAACCACATGGGC
15      R  Y  V  G  E  H  H  F  L  T  R  I  C  L  A  L  N  T  Y  G  -

CCCACGTCGAGGAGGCCGAGCGCCCCGGCGTACAGGCCGCGCGGCTGCGTTGCAGCGAC
20101 -----+-----+-----+-----+-----+ 20160
GGGTGCAGCTCCTCCGGCTCGCGGGGCCGATGTCCGGCGCCGCCGACGCAAGCTGCCTG
15      G  V  D  L  L  G  L  A  G  A  Y  L  G  R  R  S  R  E  V  S  -

TCGATGATCTCCATGGCGCGGATCTTCGGCGCGCCCGTCATGGTGCCGGCGGGGAACAGG
20161 -----+-----+-----+-----+-----+ 20220
AGCTACTAGAGGTACCGCGCCTAGAAGCCGCGCGGGCAGTACCACGGCCGCCCTTGTCC
15      E  I  I  E  M  A  R  I  K  P  A  G  T  M  T  G  A  P  F  L  -

GCGGCGATGGTGTGCAAGGCATCGGTGTCCACCCGCGCCCGGCGACGACCGTGGAGACC
20221 -----+-----+-----+-----+-----+ 20280
CGCCGCTACCACAGCTTCCGTAGCCACAGGTGGGCGCGGGCCGGCTGCTGGCACCTCTGG
15      A  A  I  T  D  F  A  D  T  D  V  R  A  R  G  V  V  T  S  V  -

AGGTGCAGCACGTGGGAGTAGCCCTCCACGTCCAGCTGGTTCGGGTACGTGAGCGTGTTC
20281 -----+-----+-----+-----+-----+ 20340
TCCACGTGTCGACCCCTCATCGGGAGGTGCAGGTGCAGCCAGCCATGCAGCTCGCACAAG
15      L  H  L  V  H  S  Y  G  E  V  D  L  Q  D  P  V  D  L  T  N  -

GGCCGGGCGATCCGTCCGATGTCGTTGCGGCAGAGGTCCACCAGCATGGTGTGCTCGGCG
20341 -----+-----+-----+-----+-----+ 20400
CCGGCCCGCTAGGCAGGCTACAGCAACGCCGTCTCCAGGTGGTCTACACACAGCCGC
15      P  R  A  I  R  G  I  D  N  R  C  L  D  V  L  M  T  H  E  A  -

ATCTCCTTGGGATCCGACCTCAGCCGGAATCCCGCGGCGATGCCGCCGTCCGCGCCGGAC
20401 -----+-----+-----+-----+-----+ 20460
TAGAGGAACCCTAGGCTGGAGTCGGCCTGAGGGCGCCGCTACGGCGGCAGGCGCGCCTG
15      I  E  K  P  D  S  R  L  R  V  G  A  A  I  G  G  D  A  G  S  -

CGCGGCACCGTGCCCGCGATCGGCCGATCGTGACCTCGCCGTCTCGATGCGTACGAAC
20461 -----+-----+-----+-----+-----+ 20520
GCGCCGTGGCACGGGCGCTAGCCGGCGTAGCACTGGAGCGGCAGGAGCTACGCATGCTTG
15      R  P  V  T  G  A  I  P  R  M  T  V  E  G  D  E  I  R  V  F  -

AGCTCGGGGCTGGCGCCGATCAGACGGTGCCCGTCGATGCCCGCCAGATACATGTACGGG
20521 -----+-----+-----+-----+-----+ 20580
TCGAGCCCCGACCGCGGCTAGTCTGCCACGGGAGCTACGGGCGGTCTATGTACATGCCC
15      L  E  P  S  A  G  I  L  R  H  G  D  I  G  A  L  Y  M  Y  P  -

GAGGCGTTCCGCCCCGCGCAGGCGCTGGTAGACGTCCGCGGGGTCCGCCGTGAGCGGATG
20581 -----+-----+-----+-----+-----+ 20640
CTCCGCAAGGCGGGCGGCTCCGCGACCATCTGCAGGCGCCCCAGCCGGCAGCTCGCCTAC
15      S  A  N  R  G  R  L  R  Q  Y  V  D  A  P  D  A  T  S  R  I  -

GAGAGCTCGTGACCGATCTGCACCTGGTAGATGTGCCGACGGCGATGTGCTTCAGACAC
20641 -----+-----+-----+-----+-----+ 20700
CTCTCGAGCACTGGCTAGACGTGGACCATCTACAGCGGCTGCCGCTACACGAAGTCTGTG
15      S  L  E  H  G  I  Q  V  Q  Y  I  D  G  V  A  I  H  K  L  C  -

CGCTCGACGTGCTTCGCGAACACTTCGGGGGCGCTGTGTCGTCGGTGACCGCGGAGGCGGGG
20701 -----+-----+-----+-----+-----+ 20760
GCGAGCTGCAGCAAGCGCTTGTGAAGCCCCCGACAGCAGCCACTGGCGCCTCCGCCCC

```

```

15      R E V D N A F V E P A S D D T V A S A P -
      AAGCCGTCTGCGGACGGATCGGGCCAGGCCTGCTCCACGTCGGCGAGGAGCCCGGTGACG
20761 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20820
      TTCGGCAGACGCCTGCCTAGCCCCGGTCCGGACGAGGTGCAGCCGCTCCTCGGGCCACTGC
15      F G D A S P D P W A Q E V D A L L G T V -
      GTCTCCGGCGCGAGGCCGGGCCAGTACGGGGACTCGTGGAGCAGCAGTTCGCATCGGCCG
20821 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20880
      CAGAGGCCGCGCTCCGGCCCCGGTTCATGCCCCCTGAGCACCTCGTTCGTCGAAGCGTAGCCGGC
15      T E P A L G P W Y P S E H L L L E C R G -
      GTGGCGAGATCGGTGACCACGCTGCCCCGGTGCAGGACCATGCGTACGTCGGCGAGGCCA
20881 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20940
      CACCGCTCTAGCCACTGGTGCAGCGGGGCCACGTCCTGGTACGCATGCAGGCCGTCCTGGT
15      T A L D T V V S G R H L V M R V D P L G -
      GGCCGGTTCTCGATGAGGTGGGGCAGGTCCTCGATGTAGCGGGCCGTGTCGTACCCGAAG
20941 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21000
      CCGGCCAAGAGCTACTCCACCCCGTCCAGGAGCTACATCGCCCGGCACAGCATGGGCTTC
15      P R N E I L H P L D E I Y R A T D Y G F -
      AACCCGAGGAACCCGAAGCGGAAGCCGGACCGGACCCCTCGGCGTCGAACATGTCCCGC
21001 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21060
      TTGGGCTCCTTGGGCTTCGCCTTCGGCCTGCGCCTGGGGAGCCGCAGCTTGTACAGGGCG
15      F G L F G F R F G S A S G E A D F M D R -
      ATGGCCCGCAGCAGCGGCCACAACCCGCCCCGCGGTACGCAGCCGCAGCCCCCTGGGGGCCG
21061 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21120
      TACGGGCGTCGTCGCCGGTGTGGGGCGGCGCCATGCGTCGGCGTCGGGGACCCCGGC
15      M A R L L P W L G G A T R L R L G Q P G -
      TCCTCCAGGAGCGCGCCGGCCCGCTCCAGGAGCAGGCCCGCAGGGCGGGTACGCCCTCG
21121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21180
      AGGAGGTCCTCGCGCGGCCGGGCGAGGTCCTCGTCCGGGGCGTCCCGCCCATGCGGGAGC
15      D E L L A G A R E L L L G R L A P V G E -
      ACGCGCACCCCGGTTCGGTGACCGAGAGCGAGAGCAGCGCGCCGAAGCCGACGAAGTGG
21181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21240
      TGCAGGTGGTGGGCCAGCCACTGGCTCTCGCTCTCGTCGCGCGGCTTCGGCTGCTTGACC
15      V R V V R D T V S L S L L A G F G V F Q -
      TGCCTGCGGTTCGCGGGCCGGGCGGCGCGGACTCCAGGAGGTAGACCTCGTCGGGGGCCG
21241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21300
      ACGGACGCCAGCGCCCGGCCCGGCGCGCCTGAGGTCTTCCATCTGGAGCAGCCCCGGC
15      H R R D R A P G A A S E L L Y V E D P G -
      AAGTGCTCGGCCAGCGCGCGGTAGGCGGGCAGGGCGCCCGTCTCCTTCACATCGAGGCGT
21301 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21360
      TTCAGAGCCGGTTCGCGCGCCATCCGCCCGTCCCGCGGGCAGAGGAAGTGTAGCTCCGCA
15      F H E A L A R Y A P L A G T E K V D L R -
      CGTGTCCGCACCCGCACCGGGCCGAGACCACGCACTGGTTCGGTCATCCTGGGTCTCTCC
21361 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21420
      GCACAGGCGTGGGCGTGGCCCCGGCTCTGGTGCCTGACCAGCCAGTAGGACCCAGGAGGG
15-<      R T R V R V P A S V V C Q D T M -
      GGATCACGTGGTGATGGCGTAGCGGTGTGCCACCTGACGGGCGGTTCAGCACCGCCCCGGTC
21421 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21480
      CCTAGTGACCACTACCGCATCGCCACACGGTGGACTGCCCGCCAGTCGTGGCGGGCCAG
14-*      * T T I A Y R H A V Q R A T L V A R D -
      GGGGCCGAGCGGTTGTTCGACGACGCGCGCGGCTTCCAGCTGACGAAGGAGCCGGTGTG
21481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21540
      CCCCAGGCTCGCCAACAGCTGCTGCGCGCGCCGAAGGTCGACTGCTTCTCGGCCACAC
14      P G S R N D V V R A A K W S V F S G T H -

```

```

GGTCACGGGGTCGAGGTCGGTGTCCACGACGATGCCGGCGTGCGCGCCGGTCCGCTCCCT
21541 -----+-----+-----+-----+-----+-----+ 21600
CCAGTGCCCCAGCTCCAGCCACAGGTGCTGTACGGCCGCACGCGCGGCCAGGCGAGGGA
14      T V P D L D T D V V I G A H A G T R E R -

GAGCCGGGCGGCGACGGCCTCGCCGATGCCCTGCCGTTCCCCCTCGGCGCCGGCCAGCAG
21601 -----+-----+-----+-----+-----+-----+ 21660
CTCGGCCCCGCGCTGCCGAGCGGCTACGGGACGGCAAGGGGGAGCCGCGGCCGGTCTGTC
14      L R A A V A E G I G Q R E G E A G A L L -

GTCCATGCGCACGGTGACGGCGTCTGCTGCCGTCGTCTGCCGGTTCGATGACGACCTGGTA
21661 -----+-----+-----+-----+-----+-----+ 21720
CAGGTACGCGTGCCACTGCCGACGCGACGGCAGCAGGACGGCCAGCTACTGCTGGACCAT
14      D M R V T V A D S G D D Q R D I V V Q Y -

GCCGAGGCAGCCGCCGACCCCGTCGAGGATCGCGGCCTCCAGCTCGGCGGGCTGGAGGGT
21721 -----+-----+-----+-----+-----+-----+ 21780
CGGCTCCGTCGGCGGGCTGGGGCAGCTCCTAGCGCCGGAGGTCGAGCCGCCGACCTCCCA
14      G L C G G V G D L I A A E L E A P Q L T -

CACGTCGCCCCAGGGGGATGCGGTCCGCGACCCGGCCGATGACCTGGATCCGCGGTCCCGG
21781 -----+-----+-----+-----+-----+-----+ 21840
GTGCAGCGGGTCCCCCTACGCCAGGCGCTGGGCCGGTACTGGACCTAGGCGCCAGGGCC
14      V D G L P I R D A V R G I V Q I R P G P -

CAGCGGCTCCCCGGGGCCCGCCGGGAGGATGCGGACCAGGTCCCCGGTTCGGTAGCGGAT
21841 -----+-----+-----+-----+-----+-----+ 21900
GTCGCCGAGGGGGCCCGGGCGGCCCTCTACGCTGGTCCAGGGGCCACGCCATCGCCTA
14      L P E G P G A P L I R V L D G T R Y R I -

CAGTGGTTTGTATGCCGTCCACCAGCATGGTGAGGACGAGTTCGCCCTCTCCCGTGTGCC
21901 -----+-----+-----+-----+-----+-----+ 21960
GTCACCAAATAACGGCAGGTGGTCTGTAACCTCTGCTCAAGCGGGAGAGGGCACAGCGG
14      L P K I G D V L M T L V L E G E G T D G -

GACCACGGCGCCGGTGTCCGGTTCGACGAGTTCGGTCAAGTAGTTGGGCTGGGCGAGGTG
21961 -----+-----+-----+-----+-----+-----+ 22020
CTGGTGCCGCGGCCACAGGCCAAGCTGCTCAAGCCAGTTCATCAACCCGACCCGCTCCAC
14      V V A G T D P E V L E T L Y N P Q A L H -

GAGCGCTCCGGTGTCCGCTCCGGTGGCGATGCACAGGGCTTCCTGGGAGCCGTAGAGCGT
22021 -----+-----+-----+-----+-----+-----+ 22080
CTCGCGAGGGCCACAGGCGAGGCCACCGCTACGTGTCCCGAAGGACCCTCGGCATCTCGCA
14      L A G T D A G T A I C L A E Q S G Y L T -

GGGCCGCACGACGGCTTGCGGCCAGAGGGTCGCCACGTTGTGCGCGAACTGCGGGGTGCA
22081 -----+-----+-----+-----+-----+-----+ 22140
CCCGGCGTGTGCGCAACGCCGGTCTCCAGCGGTGCAACAGCCGCTTGACCCCCACGT
14      P R V V A Q P W L T A V N D A F Q P T C -

GATCTCACCCAGCGTGAGGAAGAGCTTCACGGGAAGCCGGGCCAGGTCGTAGCCGTAGTG
22141 -----+-----+-----+-----+-----+-----+ 22200
CTAGAGTGGGTGCGCACTCTTCTCGAAGTGCCCTTCGGCCCGGTCCAGCATCGGCATCAC
14      I E G L T L F L K V P L R A L D Y G Y H -

CAGGGCCGCCTTGGCAAGGCTCAGGCACAGCGCCGGAGCACAGACGACCTCGACCTC
22201 -----+-----+-----+-----+-----+-----+ 22260
GTCCCGGCGGAACCGTTCCGAGTCCGTGTGCGGCCCTCGTGTCTGCTGGAGCTGGAG
14      L A A K A L S L C L A P A C V V V E V E -

CAGCTCCTCGATCAGCCGCGAGCGCTTACGGAATCCCACCCTGGGGGACTCGGGCCAGAT
22261 -----+-----+-----+-----+-----+-----+ 22320
GTCGAGGAGCTAGTCGGCGTCCGGAATGCCTTAGGGTGGGACCCCTGAGCCCGGTCTA
14      L E E I L R L A K R F G V R P S E P W I -

```



23101 AGGATCATCCGGTTGAGCAGGGCATTGACGGTCAGCTGAGCCCATACTCGCCGGCGCTG 23160  
 21 -----+-----+-----+-----+-----+-----+-----+  
 TCCTAGTAGGCCAACTCGTCCCCTAACTGCCAGTCGACTCGGGTATGGAGCGGCCGCGAC  
 L I M R N L L A N V T L Q A W V E G A S -  
 23161 TAGCGGCGGGCGACCGAGATGATCCCCGCGACCTTGTTGCTCAGCGGCCGGTCTGAAGCGC 23220  
 21 -----+-----+-----+-----+-----+-----+-----+  
 ATCGCCGCCCGCTGGCTCTACTAGGGGCGCTGGAACAACGAGTCGCCGGCCAGCTTCGCG  
 Y R R A V S I I G A V K N S L P R D F R -  
 23221 AGATAACCGACTCCGGCACGCTCGATGAAGGTCTGCATGAGGCTGGCCGTGCCGAATCCG 23280  
 21 -----+-----+-----+-----+-----+-----+-----+  
 TCTATTGGCTGAGGCCGTGCGAGCTACTTCCAGACGTAAGTCCGACCGGCACGGCTTAGGC  
 L Y G V G A R E I F T Q M L S A T G F G -  
 23281 TGCACGGGCGCCCGAAGATGATCCCGTCCGCGCGACCATCTTCGCCACGACCTCGGGC 23340  
 21 -----+-----+-----+-----+-----+-----+-----+  
 ACGTGCCCGCGGCGCTTCTACTAGGGCAGGCGGCGCTGGTAGAAGCGGTGCTGGAGCCCG  
 H V P A A F I I G D A A V M K A V V E P -  
 23341 ACCCCGTCGGCCAGGGTGCAGGCCACCGGCCTGTCTGTTGCAGTCCCCGCAGGGCCCCGCAC 23400  
 21 -----+-----+-----+-----+-----+-----+-----+  
 V G D A L T C A V P R D N C D G C P G C -  
 23401 CGCTCCATCCTGATCGAGCGCAGGTTCGACGGCCTCGAAGTCGACGCCGCGGTTCTCTGCT 23460  
 21 -----+-----+-----+-----+-----+-----+-----+  
 GCGAGGTAGGACTAGCTCGCGTCCAGCTGCCGGAGCTTCAGCTGCGGCGCCAAGAGACGA  
 R E M R I S R L D V A E F D V G R N E A -  
 23461 ACGCGTGCCGCGTGCCGCGAGTACGTGCGCGGTGTTGCCGTACGTTCCGAACCGTTGATC 23520  
 21 -----+-----+-----+-----+-----+-----+-----+  
 TGCGCACGGCGCACGGCGTTCATGCAGCGGCCACAACGGCAGTGCAAGGCTTGGAACCTAG  
 V R A A H R L V D A T N G D R E S G N I -  
 23521 GCGAGGATCTTGAGTTGTGCGCTCACGAGGGGCCTCCTTGGTGAGTCAGGTGCGCTCGGC 23580  
 13-\* -----+-----+-----+-----+-----+-----+-----+  
 21-< CGCTCCTAGAACTCAACACGCGAGTGCTCCCCGAGGAACCACTCAGTCCACGCGAGCCG  
 \* T R E A -  
 23581 GGTCCGCTCGGGGAACTGTCTGGCCGCCGCTGGTCCGGGAGCCGACGGGCCGGCTCGGC 23640  
 13 -----+-----+-----+-----+-----+-----+-----+  
 CCAGCCGAGCCCCCTTGACAGACCGGCGGCGACAGGCCCTCGGCGTCCCGGCCGAGCCG  
 T P E P S S D P R R Q D P L R L A P E A -  
 23641 GGGGGCGGGAGGAAGACCGCCCCGCGGCGGGCCGCCACGCTCGCCGAACCGGATGAGGGG 23700  
 13 -----+-----+-----+-----+-----+-----+-----+  
 CCCCCGCCCTCCTTCTGGCGGGGCGCGCCCGGCGGTGCGAGCGGCTTGGCCTACTCCCC  
 P A P P L G G R P P G G R E G F R I L P -  
 23701 CTTCTCGACGAGATAGAAGCTGATGGTCGCCAGCACGACGCTGATCGAGATCGTGAAGAG 23760  
 13 -----+-----+-----+-----+-----+-----+-----+  
 GAAGAGCTGCTCTATCTTCGACTACCAGCGGTGCTGCTGCGACTAGCTCTAGCACTTCTC  
 K E V L Y F S I T A L V V S I S I T F L -  
 23761 GAACAGTTCACGAAACCCATGTACCCCCGGAATTCCGGCGTTGGCACGGGAGACTTGCC 23820  
 13 -----+-----+-----+-----+-----+-----+-----+  
 CTTGTCAAGGGTCTTGGGGTACAGTGGGGCCTTAAGGCCGCAACCGTGCCCTCTGAACGG  
 F L E W F G M D G R F E P T P V P S K G -  
 23821 GAAGATGCTGCCGTTCTCTGAGCCAGAGGTTGATCACGATCTCGTGCCAGAGGTAGACGCC 23880  
 13 -----+-----+-----+-----+-----+-----+-----+  
 CTTCTACGACGGCAAGGACTCGGTCTCCAAGTGTAGAGCACGGTCTCCATCTGCGG  
 F I S G N R L W L N I V I E H W L Y V G -  
 GAGGGAGATCTGGCCGAGGAAGAGGATCGGCTTGCTGGTGAAGAGCGCGTCCGAGAACCG

```

23881 -----+-----+-----+-----+-----+-----+-----+ 23940
CTCCCTCTAGACCGGCTCCTTCTCTAGCCGAACGACCACTTCTCGCGCAGGCTCTTGGC
13      L S I Q G L F L I P K S T F L A D S F R -

GGACTCGGCGCCGGGGACCGTCATCGGTGCCAGGAGCAGCAGGGTGAAGGAGGTCAGGAT
23941 -----+-----+-----+-----+-----+-----+-----+ 24000
CCTGAGCCGCGGGCCCTGGCAGTAGCCACGGTCCTCGTCCCACTTCTCCAGTCCTA
13      S E A G P V T M P A L L L L T F S T L I -

GAAGTGGTTCGACGAGCTCCTGGGCCAGGGCCGCGTTGTGCGCCCATGCCCGGGATGCCGAT
24001 -----+-----+-----+-----+-----+-----+-----+ 24060
CTTCACCAGCTGCTCGAGGACCCGGTCCCGGCGCAACAGCGGGTACGGGCCCTACGGCTA
13      F H D V L E Q A L A A N D G M G P I G I -

GGGCTTGGTGGCGTAGAGGAGGTACAGCGGGATGAGCGGGACCCAGCAGATCAGCGGGCG
24061 -----+-----+-----+-----+-----+-----+-----+ 24120
CCCGAACCACCGCATCTCCTCCATGTGCGCCCTACTCGCCCTGGGTCTGTCTAGTCGCCCCG
13      P K T A Y L L Y L P I L P V W C I L P R -

CCGGATCACGAAACGGTAGAAGCCCGGGGTCCCTGGCGTCGCCTCGGCGTACGCGGAGTA
24121 -----+-----+-----+-----+-----+-----+-----+ 24180
GGCCTAGTGCTTTGCCATCTTCGGGCCCCAGGGACCGCAGCGGAGCCGCATGCGCCTCAT
13      R I V F R Y F G P T G P T A E A Y A S Y -

GATGGCCAGTGCCATGCCCGCGGCGAAGCAGCCGGCGTAGTAGGGCGGCCAGTACCACTG
24181 -----+-----+-----+-----+-----+-----+-----+ 24240
CTACCGGTACGCGTACGGGCGCGCTTCGTGCGCCGCATCATCCCGCCGGTCATGGTGAC
13      I A L A M G A A F C G A Y Y P P W Y W Q -

CATCGTCGCGCCGGTGGAGGGGAGGTTGGTGTACGTGACCCAGCCGATGGCCATGACTTC
24241 -----+-----+-----+-----+-----+-----+-----+ 24300
GTAGCAGCGCGGCCACCTCCCCTCCAACCACATGCACCTGGGTGCGGTACCGGTACTGAAG
13      M T A G T S P L N T Y T V W G I A M V E -

CAGCGCGGCCAGCGGCAGCAGGAGGCGGCGTGCCTTCTGCCCGGGAGTGCTGCCGCCCCG
24301 -----+-----+-----+-----+-----+-----+-----+ 24360
GTCGCGCCGGTCGCGCTCGTCTCCGCGCACGGAAGACGGGCGCCTCACGACGGCGGGGC
13      L A A L P L L L R R A K Q G P T S G G R -

CGCGAGCCGGTGGCCGATCCAGGCGATCAGCGGCAGGGCGAGGTAGAACGTGAACCTCGGC
24361 -----+-----+-----+-----+-----+-----+-----+ 24420
GCGCTCGGCCACCGGCTAGGTCCGCTAGTCGCGCTCCCGCTCCATCTTGCACTTGAGCCG
13      A L R H G I W A I L P L A L Y F T F E A -

GGGGACCGTCCAGGTGGGCTCGATGCCGTGCATCGGCTGGCCCTCGGGCAGATAGAAGTG
24421 -----+-----+-----+-----+-----+-----+-----+ 24480
CCCCTGGCAGGTCCACCCGAGCTACGGCACGTAGCCGACCGGGAGCCCGTCTATCTTCAC
13      P V T W T P E I G H M P Q G E P L Y F H -

CATGAGCAGCACGGGCCGAGGACGTGCTGACGCTGTGATCTCGAACCAGTTGTAGCC
24481 -----+-----+-----+-----+-----+-----+-----+ 24540
GTACTCGTCTGCGCCGCGTCTGTCAGCGACTGCGACAGCTAGAGCTTGGTCAACATCGG
13      M L L V P R L V D S V S D I E F W N Y G -

GGGGATTGCGAAGACGAGCAACAGGTAGTAGGCGGGCAGGATGCGCAGGGCCCCGGCGTTT
24541 -----+-----+-----+-----+-----+-----+-----+ 24600
CCCCTAACGCTTCTGCTCGTTGTCCATCATCCGCCCCGTCTACGCGTCCCGGGCCGCAAA
13      P I A F V L L L Y Y A P L I R L A R R K -

GAGGAACCGTCCGGTGGCGGGCGCTTCGTCCCACTGATGGTGACGCGGGCGTAGGGCTT
24601 -----+-----+-----+-----+-----+-----+-----+ 24660
CTCCTTGGCAGGCCACCGCCCGGCGAAGCAGGGTGACTACCACTGCGCCCGCATCCCGAA
13      L F R G T A P R K T G S I T V R A Y P K -

GTACAGCATCATTCGGACAGAGCGAAGAAGGGGAAGGCATACCCCCAGACCGTCCGCG
24661 -----+-----+-----+-----+-----+-----+-----+ 24720

```

13-< CATGTCGTAGTAAGGCCTGTCTCGCTTCTTCCCCCTTCCGTATGGGGGTCTGGCAGGCGC  
 Y L M M G S L A F F P S P M -  
 24721 AGGACGCCCCAGAACGGTTTGGCCGGCTCACCGACGAAGCTGCCCCACTCCGGCCTGGAAG  
 -----+-----+-----+-----+-----+-----+ 24780  
 TCCTGCGGGGTCTTGCCAAACGGGCCGAGTGGCTGCTTCGACGGGTGAGGCCGACCTTC  
 GCGACGTGGTAGACGACCACACCCAGCGCGAGGACACCTCGCAGTCCCTCGAACTTCGGT  
 24781 -----+-----+-----+-----+-----+-----+ 24840  
 CGCTGCACCATCTGCTGGTGTGGGTGCGGCTCCTGTGGAGCGTCAGGGAGCTTGAAGCCA  
 ATTCGCTTGCTTTTGGCCACCTGCGTCGCGAAGGACGTCCCCATGGAACAGTCCCCCT  
 24841 -----+-----+-----+-----+-----+-----+ 24900  
 TAAGCGAACGAAAAACGCGGTGGACGCGAGCGCTTCCTGCAGGGGTACCTTGTGAGGGGA  
 TTCCCTTGGCACTTGCTCGTTGACTTCCCGAAATAGTCGGGTCTGCGGAGTGTGAGCCGC  
 24901 -----+-----+-----+-----+-----+-----+ 24960  
 AAGGGAACCGTGAACGAGCAACTGAAGGGCTTTATCAGCCCAGACGCCTCACACTCGGCG  
 ATCTCCAATCGTGCTGTTCCGGTGCTCAGGACGACTTGTTCGGCCTGAGTGGGAAGGCA  
 24961 -----+-----+-----+-----+-----+-----+ 25020  
 TAGAGGTTAGCACGACAAGGCCACGAGTCCTGCTGAACAAAGCCGACTCACCTTCCGT  
 12-\* \* S S K N R G S H S P -  
 GCCACCCCCGCGCCCCGCCTCGGCCAGACCGGGGGCCGAGGAGTCCCGTTCGAGAGGA  
 25021 -----+-----+-----+-----+-----+-----+ 25080  
 CGGTGGGGGCGGCGGGGCGGAGCCGGTCTGGCCCCCGGCTCCTCAGGGCAAGGCTCTCCT  
 12 L W G R R G A E A L G P A S S D R E S L -  
 TCGGAGTGATCTCCGGCGGCCAGGCGATGCCACCTCCGGATCCAGCGGATTCAAGCCAT  
 25081 -----+-----+-----+-----+-----+-----+ 25140  
 AGCCTCACTAGAGGCCGCGGTCCGCTACGGGTGGAGGCCTAGGTGCGCTAAGTTCGGTA  
 12 I P T I E P P W A I G V E P D L P N L G -  
 GTTCGAGCCGGGGTCTGTAGGCCCGGAGCACAGGTAGACGATCACCGCCTCGTCGCTCA  
 25141 -----+-----+-----+-----+-----+-----+ 25200  
 CAAGCTCGGCCCCCAGCATCCGGCGGCTCGTGTCCATCTGCTAGTGGCGGAGCAGCGAGT  
 12 H E L R P D Y A A S C L Y V I V A E D S -  
 GCGTGAGGAATCCGAAGCCCAGCCCCGCGGAGACGTACAGCGCCCGTCCGTTCTCCTCGC  
 25201 -----+-----+-----+-----+-----+-----+ 25260  
 CGCACTCCTTAGGCTTCGGGTGCGGGCGCCTCTGCATGTCGCGGGCAGGCAAGAGGAGCG  
 12 L T L F G F G L G A S V Y L A R G N E E -  
 CGAGCTCCACGGTCCGCCAGCCGCGCAAGGTGGGCGACCCACCCGGATGTCGACCACGG  
 25261 -----+-----+-----+-----+-----+-----+ 25320  
 GCTCGAGGTGCCAGGCGGTGCGGCGCTTCCACCCGCTGGGGTGGGCCTACAGCTGGTGCC  
 12 G L E V T R W G G F T P S G V R I D V V -  
 CGCCGAACACGCTGCCGCGCAGGCAGCTGAAGTACTTGGCCTGGCCGGGTACGCCCCCGG  
 25321 -----+-----+-----+-----+-----+-----+ 25380  
 GCGGCTTGTCGACGGCGCGTCCGTCGACTTCATGAACCGGACCGGCCCATGCGGGGGCC  
 12 A G F V S G R L C S F Y K A Q G P V G G -  
 CGAAGTGATGCCCCGAGCACCCCGTGGGAGGAGATCGCGCAGTTCGCCTGCCGAGGT  
 25381 -----+-----+-----+-----+-----+-----+ 25440  
 GCTTCACCTACGGGGCGTCGTGGGGCACCTCCTCTAGCGCGTCAAGCGGACGGCGTCCA  
 12 A F H I G R L V G H S S I A C N A Q R L -  
 CGAAGGAGTGGCCTACGGTGCGGCGGAAGGGCTCGCCCTGGAACCACTCGCGAAACGAGC  
 25441 -----+-----+-----+-----+-----+-----+ 25500  
 GCTTCCTCACCGGATGCCACGCCGCTTCCCGAGCGGGACCTTGGTGAGCGCTTTGCTCG  
 12 D F S H G V T R R F P E G Q F W E R F S -  
 CCCGTTGCTACGGAAGACCTGCTTCTCCTCCGTCCACGCTCCCGAGATCCCGATCGGCT  
 25501 -----+-----+-----+-----+-----+-----+ 25560



12 GGGCAAGCAGTGCCTTCTGGACGAAGAGGAGGCAGGTGCGAGGGCTCTAGGGCTAGCCGA  
G R E D R F V Q K E E T W A G S I G I P -

25561 TCATCGCTGGCCCCCTTCTCTCGACTTCTCTCGACGACTCGCGGGAGGCGGCCGAGGGGTC  
-----+-----+-----+-----+-----+-----+-----+-----+ 25620  
AGTAGCGACCGGGGAAGAGAGCTGAAGAGAGCTGCTGAGCGCCCTCCGCCGGCTCCCCAG  
K M -

12-< CGCCGGGCCCCGTGGGAACGCCGCAGTCTAGATGCGGCGGCACCGGGGGCAGGGGGGTGCG  
25621 -----+-----+-----+-----+-----+-----+-----+-----+ 25680  
GCGGCCCCGGGCACCCCTTGCGGCGTCAGATCTACGCCGCCGTGGCCCCCGTCCCCCACCACG

25681 GACGACGTCCGCCCCACCTCAGCACACCGGGAGATGCAGGTCCGTGACGGGCGACGTGAC  
-----+-----+-----+-----+-----+-----+-----+-----+ 25740  
CTGCTGCAGGCGGGGTGGAGTCGTGTGGCCCTCTACGTCCAGCCACTGCCCCGCTGCACTG

25741 GATGCAACGGTCCGAGGCCCGGTTGCCCGGACGACGGCCACAGAGCCATCGGAGCAACG  
-----+-----+-----+-----+-----+-----+-----+-----+ 25800  
CTACGTTGCCAGGCTCCGGGCCAACGGGCCCTGCTGCCGGGTGTCTCGGTAGCCTCGTTGC

25801 GAGGCGGACCGCAGATGACCAAGCACGCCCGTGACCGCGCGGTAGTCCCTCGGCGCAGGGA  
-----+-----+-----+-----+-----+-----+-----+-----+ 25860  
CTCCGCCCTGGCGTCTACTGGTTCGTGCGGGCACTGGCGCGCCATCAGGAGCCGCGTCCCT

20-> M T K H A R D R A V V L G A G M -

25861 TGGCGGGGCTGCTCGCCGCGCGCGTCTGTCCGAGACGTACAAGGAAGTGCTGGTGATCG  
-----+-----+-----+-----+-----+-----+-----+-----+ 25920  
ACCGCCCCGACGAGCGGCGCGCGCAGGACAGGCTCTGCATGTTTCCTTACGACCACTAGC  
A G L L A A R V L S E T Y K E V L V I D -

25921 ACCGGGACCGGTTGGGCGGCACGGAGCAGCGCCGCGGTGTCCCGCACGGACGCCACGCCC  
-----+-----+-----+-----+-----+-----+-----+-----+ 25980  
TGGCCCTGGCCAAACCGCCGTGCCTCGTCCGCGGCCACAGGGCGTGCCCTGCGGTGCGGG  
R D R L G G T E Q R R G V P H G R H A H -

25981 ATGCGCTGCTGGCCAAGGGACAGCAGATCCTCAACGAACTCTTCCCCGGA CTGACACCG  
-----+-----+-----+-----+-----+-----+-----+-----+ 26040  
TACGCGACGACCGGTTCCCTGTCGTCTAGGAGTTGCTTGAGAAGGGGCCCTGAGCTGTGGC  
A L L A K G Q Q I L N E L F P G L D T E -

26041 AACTCACCTCGGCCCGGAATCCCCGCCGGGGACATCGCCGGGAACCTGCGGTGGTACTTCA  
-----+-----+-----+-----+-----+-----+-----+-----+ 26100  
TTGAGTGGAGCCGGCCTTAGGGGCGGCCCCCTGTAGCGGCCCTTGGACGCCACCATGAAGT  
L T S A G I P A G D I A G N L R W Y F N -

26101 ACGGCCGCCGGCTCCAGCCCTTCGACACCGGGCTGATCAGCGTCTCGGCGACGAGGCCCG  
-----+-----+-----+-----+-----+-----+-----+-----+ 26160  
TGCCGGCGGCGCGAGGTCGGGAAGCTGTGGCCCCACTAGTCGCAGAGCCGCTGCTCCGGGC  
G R R L Q P F D T G L I S V S A T R P E -

26161 AGCTGGAGTCCCACGTGCGCGCACGGGTGCGCCGCGCTGCCACAGGTGAAGATCATGGACG  
-----+-----+-----+-----+-----+-----+-----+-----+ 26220  
TCGACCTCAGGGTGACGCGCGTGCCAGCGGCGCGACGGTGTCCACTTCTAGTACCTGC  
L E S H V R A R V A A L P Q V K I M D G -

26221 GGTGCGTGATCCGGGGCCCTGACCGCCTCGGCCGACCGCAGCCGCGTACCGGTGTGAGG  
-----+-----+-----+-----+-----+-----+-----+-----+ 26280  
CCACGCACTAGGCCCGGACTGGCGGAGCCGGCTGGCGTGGCGCAGTGGCCACAGCTCC  
C V I R G L T A S A D R S R V T G V E V -

26281 TGGTCGACGAGTCGGGTACGGACACCCCCGACGCGCCTGGAGGCCGACCTCGTCTGTCAGC  
-----+-----+-----+-----+-----+-----+-----+-----+ 26340  
ACCAGCTGCTCAGCCCATGCCTGTGGGGCTGCGCGGACCTCCGGCTGGAGCAGCAGCTGC  
V D E S G T D T P T R L E A D L V V D V -

TCACGGGGCGCGGCTCGCGGACTCCCGCCTGGCTGGAGGAGTTCCGATACGAGCGGCCCCG

26341 -----+-----+-----+-----+-----+-----+-----+ 26400  
 AGTGCCCCGCGCCGAGCGCCTGAGGGCGGACCGACCTCCTCAAGCCTATGCTCGCCGGGC  
 20 T G R G S R T P A W L E E F G Y E R P A -  
 CGGAGGACCGCTTCAAGATCGATCTGGCGTACACCACGCGCCACTTCAAGCTCAAGGAAG  
 26401 -----+-----+-----+-----+-----+-----+ 26460  
 GCCTCCTGGCGAAGTTCTAGCTAGACCGCATGTGGTGC GCGGTGAAGTTCGAGTTCCTTC  
 20 E D R F K I D L A Y T T R H F K L K E D -  
 ACCCCTACGGCACGGACCTGTGATCAACCCGGTGGCATCGCCGAGCAACCCGCGCGGGC  
 26461 -----+-----+-----+-----+-----+-----+ 26520  
 TGGGGATGCCGTGCCTGGACAGCTAGTTGGGCCACCGTAGCGGCTCGTTGGGCGCGCCGC  
 20 P Y G T D L S I N P V A S P S N P R G A -  
 CGTTCTTCCCCCGGCTCGCGGACGGCAGCTCCCAGCTCTCCCTCACCGGAATCCTCGGCG  
 26521 -----+-----+-----+-----+-----+-----+ 26580  
 GCAAGAAGGGGGCCGAGCGCCTGCCGTGAGGGTCGAGAGGGAGTGGCCTTAGGAGCCGC  
 20 F F P R L A D G S S Q L S L T G I L G D -  
 ACCACCCGCCCCACCGACGACGAGGGCTTCTGGCGTTCCGCAAGTCGCTTGCCGCGCCGG  
 26581 -----+-----+-----+-----+-----+-----+ 26640  
 TGGTGGGCGGGTGGCTGCTGCTCCCGAAGGACCGCAAGCGGTTGAGCGAACGGCGCGGCC  
 20 H P P T D D E G F L A F A K S L A A P E -  
 AGATCTACCGGGCCGTCCGCGATGCCGAACCTCTCGACGAACCGGTCACCTTCCGCTTCC  
 26641 -----+-----+-----+-----+-----+-----+ 26700  
 TCTAGATGGCCCGGCAGGCGCTACGGCTTGGAGAGCTGCTTGGCCAGTGAAGGCGAAGG  
 20 I Y R A V R D A E P L D E P V T F R F P -  
 CGGCGAGCGTCCGCGCCGTTACGAGAGGCTGCGCCGTTTCCCCGGCGGGTTCTCGTCA  
 26701 -----+-----+-----+-----+-----+-----+ 26760  
 GCCGCTCGCAGGCGGCGGCAATGCTCTCCGACGCGGCAAAGGGCGCCCAAGGAGCAGT  
 20 A S V R R R Y E R L R R F P G G F L V M -  
 TGGGCGACGGCGTGTGCAGCTTCAACCCCGTCTACGGCCAGGGCATGACGGTCGCCGCCC  
 26761 -----+-----+-----+-----+-----+-----+ 26820  
 ACCCGCTGCCGCACACGTGCAAGTTGGGGCAGATGCCGGTCCCGTACTGCCAGCGGCGGG  
 20 G D G V C S F N P V Y G Q G M T V A A L -  
 TGGAGGCCGTGGCGCTGCGGGACCACTTGC GCGACGCCCCGGACCCCGACGCCCTGCGCT  
 26821 -----+-----+-----+-----+-----+-----+ 26880  
 ACCTCCGGCACCGCGACGCCCTGGTGAACGCGCTGCGGGGCTGGGGCTGCGGGACGCGA  
 20 E A V A L R D H L R D A P D P D A L R F -  
 TCTTCCGGCGTATCTCCACGGTCATCGACGTTCCGTGGGACATCGCCGCCGGAGCGGATC  
 26881 -----+-----+-----+-----+-----+-----+ 26940  
 AGAAGGCCGCATAGAGGTGCCAGTAGCTGCAAGGCACCCTGTAGCGGCGGCCTCGCCTAG  
 20 F R R I S T V I D V P W D I A A G A D L -  
 TGAACCTCCCCGGGGTGGAGGGCCCCCGCACCATGAAGGTGAAGATGGCCAACGCCTACA  
 26941 -----+-----+-----+-----+-----+-----+ 27000  
 ACTTGAAGGGGCCCCACCTCCCGGGGGCGTGGTACTTCCACTTCTACCGGTTGCGGATGT  
 20 N F P G V E G P R T M K V K M A N A Y M -  
 TGGCCCGCCTGCACGCAGCGGCAGCCGTGACGCGCGGTTGACCGGGGCGTTCTTCCGGG  
 27001 -----+-----+-----+-----+-----+-----+ 27060  
 ACCGGGCGGACGTGCGTCCCGTCCGCGAGCTGCCGCGCCACTGGCCCCGCAAGAAGGCCC  
 20 A R L H A A A A V D G A V T G A F F R V -  
 TGGCCGGGCTGGTGGACCCCCCGCAGGCCCTGATGCGCCCCCTCCCTCGCCCTGCGGGTCA  
 27061 -----+-----+-----+-----+-----+-----+ 27120  
 ACCGGCCCCGACCACCTGGGGGGCGTCCGGGACTACGCGGGGAGGGAGCGGGACGCCAGT  
 20 A G L V D P P Q A L M R P S L A L R V M -  
 TGCGCAACTCCTCGGCGAAGCCGTGCGTCCCTTCGGGCGCCCGCTATGACCGCGCGGCC  
 27121 -----+-----+-----+-----+-----+-----+ 27180

ACGCGTTGAGGAGCCGCTTCGGCAGCCAGGGAAGCCCGCGGCGGCATACTGGCGCGCCGG  
 20-\* R N S S A K P S V P S G A A V \* -

CGTCCGGGGCGGCTGCCGGGGCCAGGAGCCGACATGCGGGTGATGATCACGGTGTTCCCG  
 27181 -----+-----+-----+-----+-----+ 27240  
 19 GCAGGCCCCCGCGACGGCCCCGGTCCTCGGCTGTACGCCCCTACTAGTGCCACAAGGGC  
 M R V M I T V F P -

GCGCGGGCGCACTTCCTGCCGCTGGTGCCCTATGCCTGGGCCCCTGCAGAGCGCGGGCCAC  
 27241 -----+-----+-----+-----+-----+ 27300  
 19 CGCGCCCGCGTGAAGGACGGCGACCACGGGATACGGACCCGGGACGTCTCGCGCCCGGTG  
 A R A H F L P L V P Y A W A L Q S A G H -

GAGGTATGTGTCGTGGCGCCCCCGGGCTATCCCACCGGGGTGGCCGACCCCGACTTCCAC  
 27301 -----+-----+-----+-----+-----+ 27360  
 19 CTCCATACACAGCACCGCGGGGGCCCGATAGGGTGGCCCCACCGGTGGGGCTGAAGGTG  
 E V C V V A P P G Y P T G V A D P D F H -

GAGGCCGTCACCGCGGCCCGCCTGAAGTCGGTGACCTGCGGGCAGCCGCAGCCGCTGGCG  
 27361 -----+-----+-----+-----+-----+ 27420  
 19 CTCCGGCAGTGGCGCCGGCCGACTTCAGCCACTGGACGCCCCGTGCGCGTGGCGGACCGC  
 E A V T A A G L K S V T C G Q P Q P L A -

GTCCACGACCGCGACGACCCCGGCTACGCGGCGATGCTGCCGACCGCGGCGGAGTCGGAG  
 27421 -----+-----+-----+-----+-----+ 27480  
 19 CAGGTGCTGGCGCTGCTGGGGCCGATGCGCCGCTACGACGGCTGGCGCCGCTCAGCCTC  
 V H D R D D P G Y A A M L P T A A E S E -

CGCTACGTGGCGGCCCTCGGGATCAGCGAGAAGGAGCGCCCCACCTGGGACGTCTTCTAC  
 27481 -----+-----+-----+-----+-----+ 27540  
 19 GCGATGCACCGCCGGGAGCCCTAGTCGCTCTTCTCGCGGGTGGACCTGCAGAAGATG  
 R Y V A A L G I S E K E R P T W D V F Y -

CACTTCACCTTGCTGGCGATCCGCGACTACCATCCGCCGCGGCCGCGGACGAGCTGGAC  
 27541 -----+-----+-----+-----+-----+ 27600  
 19 GTGAAGTGGAACGACCGCTAGGCGCTGATGGTAGGCGGCGCCGGCGCCGTCTGCACCTG  
 H F T L L A I R D Y H P P R P R Q D V D -

CAGGTGATCGAGTTTCGCCCCGATCTGGCAGCCCCGATCTGGTGCTGTGGGACGCTGGTTC  
 27601 -----+-----+-----+-----+-----+ 27660  
 19 GTCCACTAGCTCAAGCGGGCCTAGACCGTTCGGGCTAGACCACGACACCCTGCGGACCAAG  
 Q V I E F A R I W Q P D L V L W D A W F -

CCCTCGGGCGCGATCGCGGCGCGGGTCAGCGGCGCCGCGCACGCGCGGGTGCTCGTAGCC  
 27661 -----+-----+-----+-----+-----+ 27720  
 19 GGGAGCCCGCGCTAGCGCCGCGCCAGTCGCGCGGCGCGTGCAGCGCCACGAGCATCGG  
 P S G A I A A R V S G A A H A R V L V A -

CCCCACTACACCGGTGGGTACCGAGCGGTTTCGCCGCCGCGGGCCCCCGCGGGGGGCC  
 27721 -----+-----+-----+-----+-----+ 27780  
 19 GGGCTGATGTGGCCGACCCAGTGGCTCGCCAAGCGGCGGCGCCCGGGCGCCGCCCCCGG  
 P D Y T G W V T E R F A A A G P A A G A -

GACCTCCTGGCCGAGACGATGCGGCCGCTGGCCGAGCGGTACGGCGTGGAGGTCGACGAC  
 27781 -----+-----+-----+-----+-----+ 27840  
 19 CTGGAGGACCGGCTCTGCTACGCCGCGACCGGCTCGCCATGCCGCACCTCCAGCTGCTG  
 D L L A E T M R P L A E R Y G V E V D D -

GATCTTCTGCTCGGACAGTGGACGGTCAATCCGTTCCCGGCGCCGATGAACCCGCCGACC  
 27841 -----+-----+-----+-----+-----+ 27900  
 19 CTAGAAGACGAGCCTGTACCTGCCAGTTAGGCAAGGGCCGCGGCTACTTGGGCGGCTGG  
 D L L L G Q W T V N P F P A P M N P P T -

CGGCTCACGAACGTTCCGGTGCGCTACGTGCCCTACACCGGTGCCAGCGTCATGCCCGCG  
 27901 -----+-----+-----+-----+-----+ 27960  
 GCCGAGTGCTTGCAAGGCCACGCGATGCACGGGATGTGGCCACGGTCGCGAGTACGGGCGC



```

1      R V T V L D K L T Y S G N P A N L T S V -
      TCGCGGCCCATCCGCGGTACACCTTCGTCCAGGGCGACACCGTCGACCCGCGCGTCGTGCG
28741 -----+-----+-----+-----+-----+-----+-----+ 28800
      AGCGCCGGGTAGGCGCCATGTGGAAGCAGGTCCCGCTGTGGCAGCTGGGCGCGCAGCAGC
1      A A H P R Y T F V Q G D T V D P R V V D -
      ACGAGGTGGTTCGCCGGCCACGACGTATCGTCCACTTCGCGGCGGAGTCGCACGTGGACC
28801 -----+-----+-----+-----+-----+-----+-----+ 28860
      TGCTCCACCAGCGGCCGGTGTCTGCAGTAGCAGGTGAAGCGCCGCCTCAGCGTGCACCTGG
1      E V V A G H D V I V H F A A E S H V D R -
      GCTCGATCGACACCGCCACCCGGTTCGTACAGACCAACGTGCTCGGGACCCAGACGCTGC
28861 -----+-----+-----+-----+-----+-----+-----+ 28920
      CGAGCTAGCTGTGGCGGTGGGCCAAGCAGTGTGTTGCACGAGCCCTGGGTCTGCGACG
1      S I D T A T R F V T T N V L G T Q T L L -
      TGGAAGCGGCTCTCCGGCACGGGGTTCGGCCGGTTCGTGCACGTGTGACCGACGAGGTCT
28921 -----+-----+-----+-----+-----+-----+-----+ 28980
      ACCTTCGCCGAGAGGCCGTGCCCCAGCCGGCCAAGCACGTGCACAGCTGGCTGCTCCAGA
1      E A A L R H G V G R F V H V S T D E V Y -
      ACGGGTCGATCGCCTCCGGCTCATGGACCGAGGACACCCCGCTCGCCCCAACGTCCCCT
28981 -----+-----+-----+-----+-----+-----+-----+ 29040
      TGCCGAGCTAGCGGAGGCCGAGTACCTGGCTCCTGTGGGGCGAGCGGGGTTCAGGGGA
1      G S I A S G S W T E D T P L A P N V P Y -
      ACGCGGCGTCGAAGGCGGGTTCGGACCTGATGGCGCTCGCCTGGCACCGCACCCGGGGCC
29041 -----+-----+-----+-----+-----+-----+-----+ 29100
      TGCGCCGCGAGCTTCGCCCAAGCCTGGACTACCGCGAGCGGACCGTGGCGTGGGCCCCGG
1      A A S K A G S D L M A L A W H R T R G L -
      TGGACGTGTCGTACCCCGGTGCACCAACAACCTACGGTCCCTACCAGTACCCCGAGAAGG
29101 -----+-----+-----+-----+-----+-----+-----+ 29160
      ACCTGCAGCAGCAGTGGGCCACGTGGTTGTTGATGCCAGGGATGGTCATGGGGCTCTTCC
1      D V V V T R C T N N Y G P Y Q Y P E K V -
      TGATCCCGCTCTTCGTACCAACATCCTCGACGGCTTGCGGGTGCCCTGTACGGGGACG
29161 -----+-----+-----+-----+-----+-----+-----+ 29220
      ACTAGGGCGAGAAGCAGTGGTTGTAGGAGCTGCCGAACGCCACGGGGACATGCCCTGTC
1      I P L F V T N I L D G L R V P L Y G D G -
      GCGCCACCGCCGGGACTGGCTGCACGTGTCCGACCCTGCCGGGCCATCCAGATGGTCA
29221 -----+-----+-----+-----+-----+-----+-----+ 29280
      CGCGGGTGGCGGCCCTGACCGACGTGCACAGGCTGGTGACGGCCCGGTAGGTCTACCAGT
1      A H R R D W L H V S D H C R A I Q M V M -
      TGAACTCCGGCCGGGCGGGGAGGTCTACCACATCGGCGGCGGCACCGAACTCTCCAACG
29281 -----+-----+-----+-----+-----+-----+-----+ 29340
      ACTTGAGGCCGGCCCGGCCCTCCAGATGGTGTAGCCGCCGCGTGGCTTGAGAGGTTGC
1      N S G R A G E V Y H I G G G T E L S N E -
      AGGAACTCACCGGCCTGTTGCTCACGGCGTGCGGCACCGACTGGTCCTGCGTGGACCGGG
29341 -----+-----+-----+-----+-----+-----+-----+ 29400
      TCCTTGAGTGGCCGGAACAGAGTGCCGCACGCCGTGGCTGACCAGGACGCACCTGGCCC
1      E L T G L L L T A C G T D W S C V D R V -
      TGGCCGACCGGCAGGGGCACGACCGCCGCTACTCGCTCGACATCACGAAGATCCGGCAGG
29401 -----+-----+-----+-----+-----+-----+-----+ 29460
      ACCGGCTGGCCGTCCCCGTGCTGGCGGCGATGAGCGAGCTGTAGTGCTTCTAGGCCGTCC
1      A D R Q G H D R R Y S L D I T K I R Q E -
      AACTGGGCTACGAGCCCCTGGTGCCTTCGAGGACGGCCTGGCCGCGACGGTGAAGTGGT
29461 -----+-----+-----+-----+-----+-----+-----+ 29520
      TTGACCCGATGCTCGGGGACCAGCGGAAGCTCCTGCCGGACCGGCGCTGCCACTTCACCA
1      L G Y E P L V A F E D G L A A T V K W Y -

```

29521 ACCACGAGAACCGTTTCGTGGTGGCAGCCGCTGAAGGAAGCGGCCGCTCCTGGACGCCG 29580  
 1 -----+-----+-----+-----+-----+-----+-----+  
 TGGTGCTCTTGGCAAGCACCACCGTCGGCGACTTCCTTCGCCGCGCCGAGGACCTGCGGC  
 H E N R S W W Q P L K E A A G L L D A V -  
 29581 TCGGCTGACGGCAGCCACCGCTAGGAACACCCCAGGAAAGGAGCCACCTCCGTGACAGCA 29640  
 -----+-----+-----+-----+-----+-----+-----+  
 AGCCGACTGCCGTCGGTGGCGATCCTTGTGGGGTCTTTCTCCTCGGTGGAGGCACTGTCGT  
 2-> M T A -  
 1-\* G \* -  
 29641 GTCAAGGAGCCGACGTCCCGCGCAGGACGGCGGGAGTGGATCGCTCTCGTCGTCTCTCC 29700  
 -----+-----+-----+-----+-----+-----+-----+  
 CAGTTCCTCGGCTGCAGGGCGCGTCTGCCGCCCTCACCTAGCGAGAGCAGCAGGAGAGG  
 2 V K E P T S R A G R R E W I A L V V L S -  
 29701 TTGCCCACGATGCTGTTGATGCTGGACATCAACGTCTCATGCTGGCCTTGCCGCAGTTG 29760  
 -----+-----+-----+-----+-----+-----+-----+  
 AACGGGTGCTACGACAACCTACGACCTGTAGTTGCAGGAGTACGACCGGAACGGCGTCAAC  
 2 L P T M L L M L D I N V L M L A L P Q L -  
 29761 AGCGAGGATCTCGGCGCGAGCAGCACGCAACAGCTGTGGATCACCGACATCTACGGATTC 29820  
 -----+-----+-----+-----+-----+-----+-----+  
 TCGCTCCTAGAGCCGCGCTCGTCGTGCGTTGTGCGACACCTAGTGGCTGTAGATGCCTAAG  
 2 S E D L G A S S T Q Q L W I T D I Y G F -  
 29821 GCGATCGCCGGCTTCTGTTGACCATGGGCACCCTCGGCGACCGGATCGGCCGCCGAGG 29880  
 -----+-----+-----+-----+-----+-----+-----+  
 CGCTAGCGGCCGAAGGACCACTGGTACCCGTTGGGAGCCGCTGGCCTAGCCGGCGGCGTCC  
 2 A I A G F L V T M G T L G D R I G R R R -  
 29881 CTCCTGCTCGGGGGCGCGGCCGTCTTCGCGGTGCTGTCCGTGCTCGCCGCGTTCTCCGAC 29940  
 -----+-----+-----+-----+-----+-----+-----+  
 GAGGACGAGCCCCCGCGCCGCGAGAAGCGCGACAGGCGAGCGCGCCGAAGAGGCTG  
 2 L L L G G A A V F A V V S V V A A F S D -  
 29941 AGCGCGGCGATGCTCGTCTGTCAGCCGCGCCGTGCTCGGCGTCCGCCGGGGCCACGGTGATG 30000  
 -----+-----+-----+-----+-----+-----+-----+  
 TCGCGCCGCTACGAGCAGCAGTCCGCGCGGCACGAGCCGAGCGGCCCGGTGCCACTAC  
 2 S A A M L V V S R A V L G V A G A T V M -  
 30001 CCCTCGACGCTCGCGCTCATCAGCAACATGTTTCGAGGACCCCAAGGAGCGGGGCACCGCC 30060  
 -----+-----+-----+-----+-----+-----+-----+  
 GGGAGCTGCGAGCGCGAGTAGTCGTTGTACAAGCTCCTGGGGTTCTCGCCCCGTGGCGG  
 2 P S T L A L I S N M F E D P K E R G T A -  
 30061 ATCGCCATGTGGGCGAGCGCCATGATGGCCGGAGTCGCCCTCGGGCCCGCGTCCGCGGC 30120  
 -----+-----+-----+-----+-----+-----+-----+  
 TAGCGGTACACCCGCTCGCGGTACTACCGGCCTCAGCGGGAGCCCGGGCGGACGCCGCCG  
 2 I A M W A S A M M A G V A L G P A V G G -  
 30121 CTGGTCTCTCGCCGCTTCTGGTGGGGATCGGTGTTCTCATCGCCGTTCCGGTGATGCTG 30180  
 -----+-----+-----+-----+-----+-----+-----+  
 GACCAGGAGCGGCGCAAGACCACCCCTAGCCACAAGGAGTAGCGGCAAGGCCACTACGAC  
 2 L V L A A F W W G S V F L I A V P V M L -  
 30181 CTGGTGGTGGTCAACGGCCCCGTGCTGCTCACCGAGTCCCGCGACCCGGACGCCGGACGG 30240  
 -----+-----+-----+-----+-----+-----+-----+  
 GACCACCACAGTGGCCGGGGCACGACGAGTGGCTCAGGGCGCTGGGCCTGCGGCTGCC  
 2 L V V V T G P V L L T E S R D P D A G R -  
 30241 CTGGACCTGCTGAGCGCGGGGCTCTCCCTCGCGACCGTGCTGCCGGTGATCTACGGACTG 30300  
 -----+-----+-----+-----+-----+-----+-----+  
 GACCTGGACGACTCGCGCCCCGAGAGGGAGCGCTGGCACGACGGCCACTAGATGCCTGAC  
 2 L D L L S A G L S L A T V L P V I Y G L -

AAGGAGCTGGCCCGGACCGGGTGGGACCCGCTCGCCGCGGCGCGGTGGTCCCTCGGCGTG  
 30301 -----+-----+-----+-----+-----+ 30360  
 TTCCTCGACCGGGCCTGGCCACCCCTGGGCGAGCGGCGGCGCCACCAGGAGCCGCAC  
 2 K E L A R T G W D P L A A G A V V L G V -  
  
 ATCTTCGGCGCGCTGTTCTGTCAGCGCCAGCGGCGGTTGGCCGACCCCATGCTGGACCTC  
 30361 -----+-----+-----+-----+-----+ 30420  
 TAGAAGCCGCGGACAAGCAGGTGCGGGTGGCGGCAACCGGCTGGGGTACGACCTGGAG  
 2 I F G A L F V Q R Q R R L A D P M L D L -  
  
 GGCCTCTTCGCGGACCGCACCCCTGCGGGCGGGTCTGACGGTCAGTCTGGTCAACGCCGTC  
 30421 -----+-----+-----+-----+-----+ 30480  
 CCGGAGAAGCGGCTGGCGTGGGACGCGGCGGCGGAGTCCAGTCCAGACAGTTGCGGCAG  
 2 G L F A D R T L R A G L T V S L V N A V -  
  
 ATCATGGGCGGGACCGGACTGATGGTCGCCCTGTACCTCCAGACGATCGCCGGTCACTCC  
 30481 -----+-----+-----+-----+-----+ 30540  
 TAGTACCGGCCCTGGCCTGACTACCAGCGGGACATGGAGGTCTGCTAGCGGCCAGTGAGG  
 2 I M G G T G L M V A L Y L Q T I A G H S -  
  
 CCGTTGGCGCGCGGGCTGTGGCTGCTGATCCCGGCCTGCATGCTCGTCTGGGCGTACAG  
 30541 -----+-----+-----+-----+-----+ 30600  
 GGCAACCGGCGGCGGCGGACACCGGAGTGGGCGGACGTACGAGCAGCACCCGCATGTC  
 2 P L A A G L W L L I P A C M L V V G V Q -  
  
 CTGTGCAACCTGCTGGCCAGCGGATGCCCCCTTCCCGGGTGTGCTGGGGGGACTGCTG  
 30601 -----+-----+-----+-----+-----+ 30660  
 GACAGCTTGGACGACCGGGTGCCTACGGGGGAAGGGCCACGACGACCCCCCTGACGAC  
 2 L S N L L A Q R M P P S R V L L G G L L -  
  
 ATCGCGCGCGTGGACAGCTCCTGATCACCCAGGTGGACACCGAGGACACCGCCCTCCTC  
 30661 -----+-----+-----+-----+-----+ 30720  
 TAGCGCGGCGAGCCTGTGAGGACTAGTGGGTCCACCTGTGGCTCCTGTGGCGGGAGGAG  
 2 I A A V G Q L L I T Q V D T E D T A L L -  
  
 ATCGCGGCCACCAACCTGATCTACTTCGGCGCCTCACCGGTGGGGCCGATCACACGCGGC  
 30721 -----+-----+-----+-----+-----+ 30780  
 TAGCGCGGCTGGTGGGACTAGATGAAGCCGCGGAGTGGCCACCCCGGCTAGTGGTGCCCCG  
 2 I A A T T L I Y F G A S P V G P I T T G -  
  
 GCGATCATGGGAGCCGCGCCCCCGGAGAAGGCGGGTGGCGCCTCGTCGCTGTCCGCCACC  
 30781 -----+-----+-----+-----+-----+ 30840  
 CGCTAGTACCCTCGGCGCGGGGGCCTCTTCCGCCACGGCGGAGCAGCGACAGGCGGTGG  
 2 A I M G A A P P E K A G A A S S L S A T -  
  
 GGCGGCGAGTTCGGAGTGGCGCTCGGCATCGCGGGCCTGGGGAGTCTGGGCACCGTCTGTG  
 30841 -----+-----+-----+-----+-----+ 30900  
 CCGCGGCTCAAGCCTCACCGGAGCCGTAGCGCCCGGACCCCTCAGACCCGTGGCAGCAC  
 2 G G E F G V A L G I A G L G S L G T V V -  
  
 TACAGCGCCGGGGTGGAGTGCCGACGCGGCGGGCCCGGACCGGACCGCGCGCAG  
 30901 -----+-----+-----+-----+-----+ 30960  
 ATGTCGCGGCCCCAGCTCCACGGCCTGCGCCGGCGGCGGCTGCGGCTGCGGCGCGTC  
 2 Y S A G V E V P D A A G P A D A D A A Q -  
  
 GAGAGCATCGCCGGCGCCCTGCACACGGCCGGTCACTGGCACCGGGCAGCGCCGACGCC  
 30961 -----+-----+-----+-----+-----+ 31020  
 CTCTCGTAGCGGCCGCGGGACGTGTGCCGGCCAGTCGACCGTGGCCCGTCCGCGGCTGCGG  
 2 E S I A G A L H T A G Q L A P G S A D A -  
  
 CTGCTGGACTCCGCGCGCGCGGCCTTACCAGCGGCGTGCAGTCCGTGCGCGCCGCTCTGC  
 31021 -----+-----+-----+-----+-----+ 31080  
 GACGACCTGAGGCGCGCGCGCGGAAGTGGTGGCGCACGTCAGGCAGCGGCGGCAGACG  
 2 L L D S A R A A F T S G V Q S V A A V C -

```

31081  GCGGTGTTCTCCCTGGCGCTCGCCGTCCTCATCGGCACCCGGCTGCGGGACATTTCCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31140
2      CGGCACAAGAGGGACCGCGAGCGGCAGGAGTAGCCGTGGGCGGACGCCCTGTAAAGGCGC
A V F S L A L A V L I G T R L R D I S A -

31141  ATGGACCACGGGCACGGCGAGGAACCGGCCGAGAACGACGCTCAACCGGCCACATGAGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31200
2-*    TACCTGGTGCCCGTGCCGCTCCTTGCCCGGCTCTTGCTGCGAGTTGGCCGGTGTACTCGC
M D H G H G E E P A E N D A Q P A T * -

31201  CACTTCCGGAGATGCAACGGCCGCGCTCGAGGTATGAGGATCACCTTCCGGGGTGACCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31260
GTGAAGGCCTCTACGTTGCCGGCGGCGAGCTCCATACTCCTAGTGGAAGGCCCCACGTGGA

31261  GCACGGCAACGGAGGCGTAGTGGAGTACTGGAACAGCACGGCGGAGACCATGCCCCGCCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31320
3->    CGTGCCGTTGCCTCCGCATCACCTCATGACCTTGTCGTGCCGCTCTGGTACGGGGCGGT
M E Y W N S T A E T M P R Q -

31321  GGAACTCGAACAGTGGAAGTGGCGCAGGCTCCAGGCCGCCATGGACCACGCCAGAAGGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31380
3      CCTTGAGCTTGTACCTTCACCGCGTCCGAGGTCCGGCGGTACCTGGTGCGGTCTTCCGA
E L E Q W K W R R L Q A A M D H A R R L -

31381  TTCGCCCTTCTGGCGGGAACGACTCCCCGAGAACATCACCTCCATGGCGGACTACGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31440
3      AAGCGGGAAGACCGCCCTTGCTGAGGGGCTCTTGTAAGTGGAGGTACCGCCTGATGCGCCG
S P F W R E R L P E N I T S M A D Y A A -

31441  GCGGGTGCTCTCCTGCGCAAGGCCGACCTCCTCGCCGCGGAAGCCGCTCTCCCCCTTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31500
3      CGCCACGGAGAGGACGCGTTCGGGCTGGAGGAGCGGCGCCTTCGGCGCAGAGGGGGAAT
R V P L L R K A D L L A A E A A S P P Y -

31501  CGGCACCTGGCCCTCGCTGGATCCGGCGCTCGGAGTGCGCCATCACCAGACCAGCGGCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31560
3      GCCGTGGACCGGGAGCGACCTAGGCCGCGAGCCTCACGCGGTAGTGGTCTGGTGCCTGTG
G T W P S L D P A L G V R H H Q T S G T -

31561  CAGCGGTAACCCCCCATCCGGACGTTGACACCGAACGCGACTGGGCCTGGTGCGTGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31620
3      GTCGCCATTGGGGGGGTAGGCCTGCAAGCTGTGGCTTGCGCTGACCCGGACCACGCACCT
S G N P P I R T F D T E R D W A W C V D -

31621  CACGTTCTGCACGGCGCTCCACAGCATGGGCGTGCGCCCGCACCACAAGGGTCTGGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31680
3      GTGCAAGACGTGCCGCGAGGTGTGCTACCCGCACGCGGCGTGGTGTTCACAGACCACCG
T F C T A L H S M G V R P H H K G L V A -

31681  GTTCGGCTACGGGCTGTTGCGCCGTTTCTGGGGCATGCACTACGGCCTCGAGCGCATGGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31740
3      CAAGCCGATGCCCCGACAAGCGGCCAAAGACCCCGTACGTGATGCCGGAGCTCGCGTACCC
F G Y G L F A G F W G M H Y G L E R M G -

31741  CGCCACGGTCATCCCGGCCGGCGGCCTCGACTCCCGCTCCCGGGTACGGCTGCTGGTCTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31800
3      GCGGTGCCAGTAGGGCCGGCCGCGGAGCTGAGGGCGAGGGCCCATGCCGACGACGACT
A T V I P A G G L D S R S R V R L L V D -

31801  CTACCAGATCGAGGTGCTCGGCCTCACACCGAGCTATGCGATGCGGCTGATCGAGACGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31860
3      GATGGTCTAGCTCCACGAGCCGAGTGTGGCTCGATACGCTACGCCGACTAGCTCTGCCG
Y Q I E V L G L T P S Y A M R L I E T A -

31861  CCGCGAGATGGGCATCGACCTCGCCCGCGAGGCTAACGTCCAGATCATCTGGCCGGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31920

```



```

3      GGCGCTCTACCCGTAGCTGGAGCGGGCGCTCCGATTGCAGGTCTAGTAGGACCGGCCCCG
      R E M G I D L A R E A N V Q I I L A G A -
31921  GGAGCCGCGCTCCGCGTTTACCACCCGCACCATCGAGGAGGCCTTCGGCGCCCGGGTCTT
      -----+-----+-----+-----+-----+-----+-----+ 31980
3      CCTCGGCGCGAGGCGCAAGTGGTGGGCGTGGTAGCTCCTCCGGAAGCCGCGGGCCCAGAA
      E P R S A F T T R T I E E A F G A R V F -
31981  CAACGCCGCGGGCACCCTGAGTTTCGGGGGGGTGTTTCATGTTTCGAGTGCACCGCCCGGCG
      -----+-----+-----+-----+-----+-----+-----+ 32040
3      GTTTCGGCGCCCGTGGTGAAGCCCCCACAAGTACAAGCTCACGTGGCGGGCCGC
      N A A G T T E F G G V F M F E C T A R R -
32041  CGAGGCCTGCCACATCATCGAACCCTCGTGCATCGAGGAGGTGCTCGACCCGGTGACGGA
      -----+-----+-----+-----+-----+-----+-----+ 32100
3      GCTCCGGACGGTGTAGTAGCTTGGGAGCAGTAGCTCCTCCACGAGCTGGGCCACTGCCT
      E A C H I I E P S C I E E V L D P V T E -
32101  ACAGCCCGTTCGGCTACGGCGAGGAGGGCGTCCGAGTCACCACCGGGCTGAACCGTGAGGG
      -----+-----+-----+-----+-----+-----+-----+ 32160
3      TGTGGGCGAGCCGATGCCGCTCCTCCCGCAGGCTCAGTGGTGGCCCCGACTTGGCACTCCC
      Q P V G Y G E E G V R V T T G L N R E G -
32161  GATGCAGCTCTTCCGGCACTGGACCGAGGACGTCGTGGTCAAGCGGCCCCACACCGAGTG
      -----+-----+-----+-----+-----+-----+-----+ 32220
3      CTACGTCGAGAAGGCCGTGACCTGGCTCCTGCAGCACCAGTTCGCCGGGGTGTGGCTCAC
      M Q L F R H W T E D V V V K R P H T E C -
32221  CGGCTGCGGCGGACGTTGGACTTCTACGACGGCGGCATCCTTCGGCGCGTGGACGACAT
      -----+-----+-----+-----+-----+-----+-----+ 32280
3      GCCGACGCCGCGCTGCACCCTGAAGATGCTGCCGCCGTAGGAAGCCGCGACCTGCTGTA
      G C G R T W D F Y D G G I L R R V D D M -
32281  GCGCAAGATACGCGGGGTCTCGATCACCCCGGTGATGATCGAGGATGTGCTGCGCGGCTT
      -----+-----+-----+-----+-----+-----+-----+ 32340
3      CGCGTTCTATGCGCCCCAGAGCTAGTGGGGCCACTACTAGCTCCTACACGACGCGCCGAA
      R K I R G V S I T P V M I E D V L R G F -
32341  CGACGAGGTGAACGAGTTCCACTCGTCCATCCGACCGTCCGCGGACTCGATACGATCCA
      -----+-----+-----+-----+-----+-----+-----+ 32400
3      GCTGCTCCACTTGCTCAAGGTGAGCAGGTAGGCCTGGCAGGCGCCTGAGCTATGCTAGGT
      D E V N E F H S S I R T V R G L D T I H -
32401  CGTCAAGGTCGAGGCGGGGAGACATCTCGGGTGAAGCGGCCGAGAGCCTGTGCGGCCGCAT
      -----+-----+-----+-----+-----+-----+-----+ 32460
3      GCAGTTCAGCTCCGCCCTCTGTAGAGCCCACTCCGCCGGCTCTCGGACACGCGGGCGTA
      V K V E A G D I S G E A A E S L C G R I -
32461  CACCGAGGAGTTCAAGCGTGAGATAGGCATACGGCCCCAGGTGGAGCTGACCCCCGCGGG
      -----+-----+-----+-----+-----+-----+-----+ 32520
3      GTGGCTCCTCAAGTTCGCACTCTATCCGTATGCCGGGGTCCACCTCGACTGGGGGCGCCC
      T E E F K R E I G I R P Q V E L T P A G -
32521  CAGCCTCCCCCGATCGAAGTGGAAGGCGGCACGACTTCATGACGAGCGGAACTCGCCCC
      -----+-----+-----+-----+-----+-----+-----+ 32580
3      GTCGGAGGGGGTAGCTTACCTTCCGCCGTGCTGAAGTACTGCTCGCGCTTGAGCGGGG
      S L P R S K W K A A R L H D E R E L A P -
32581  TCAGGCCTGAGCAGGTGGAGCAGCTCCTGGTGAAGTACCGGAGCCTGGGCCTGCTGGAGC
      -----+-----+-----+-----+-----+-----+-----+ 32640
3-*  AGTCCGGAAGTCCACCTCGTCGAGGACCACTCGATGGCCTCGGACCCGGACGACCTCG
      Q A * -
32641  AGAGCTGCGCGGTCCCGGCCGTGCTCGCCGCGGTACAGGGCCGCCCGTGCAGAACTCCGTA
      -----+-----+-----+-----+-----+-----+-----+ 32700
      TCTCGACGCGCCAGGGCCGGCACGAGCGGCGCCAGTCCCGGCGGGCACGCCTTGAGGCAT

```

TCGCCCTGGACGGCCAGGGCGTGGAGTTCGAGTACTACCGGGGCACGACGACGACCTCG  
32701 -----+-----+-----+-----+-----+ 32760  
AGCGGGACCTGCCGGTCCCGCACCTCAAGCTCATGATGCCCCCGTGCTGCTGTCGGAGC

TGGCCTGAACCCACCCCCGGTCCGCCGGGTGACAGCAAAGGGAGACCGGTGCCCCACGGT  
32761 -----+-----+-----+-----+-----+ 32820  
ACCGGACTTGGGTGGGGGCCAGGCGGCCAGTCTGCTTTCCTCTGGCCACGGGGTGCCA  
4-> M P H G -

GCAGAGCGCGAAGCGAGCCCGGCCGAGGAGAGCGCCGGCACCCGGCCGCTGACCGGCGAG  
32821 -----+-----+-----+-----+-----+ 32880  
CGTCTCGCGCTTCGCTCGGGCCGGCTCCTCTCGCGGCCGTGGGCCGGCGACTGGCCGCTC  
4 A E R E A S P A E E S A G T R P L T G E -

GAGTATCTGGAGAGCCTGCGGGACGCGCGGGAGGTGTACCTCGACGGCAGCCGCGTCAAG  
32881 -----+-----+-----+-----+-----+ 32940  
CTCATAGACCTCTCGGACGCCCTGCGCGCCCTCCACATGGAGCTGCCGTGCGGCGAGTTCT  
4 E Y L E S L R D A R E V Y L D G S R V K -

GACGTCACCGCGCATCCCGCGTTCACAACCCGGCCCGGATGACGGCCCCGGCTGTACGAC  
32941 -----+-----+-----+-----+-----+ 33000  
CTGCAGTGGCGCGTAGGGCGCAAGGTGTTGGGCCGGGCTACTGCCGGGCCGACATGCTG  
4 D V T A H P A F H N P A R M T A R L Y D -

AGCCTGCACGACCCCGCCAGAAAGCGGTCTTGACGGCGCCACCGATGCCGGTGACGGT  
33001 -----+-----+-----+-----+-----+ 33060  
TCGGACGTGCTGGGGCGGGTCTTTCGCCAGGACTGCCGCGGGTGCTACGGCCACTGCCA  
4 S L H D P A Q K A V L T A P T D A G D G -

TTCACCCACCGCTTCTTACCGCACCGCGCAGCGTCGACGACCTGGTCAAGGACCAGGCC  
33061 -----+-----+-----+-----+-----+ 33120  
AAGTGGGTGGCGAAGAAGTGGCGTGGCGCGTCGACGCTGCTGGACCAGTTCTTGGTCCGG  
4 F T H R F F T A P R S V D D L V K D Q A -

GCCATCGCATCCTGGGCGCGCAAGAGCTACGGCTGGATGGGGCGCAGCCCCGACTACAAG  
33121 -----+-----+-----+-----+-----+ 33180  
CGGTAGCGTAGGACCCGCGCGTTCGATGCCGACCTACCCCGCGTCGGGGCTGATGTTCT  
4 A I A S W A R K S Y G W M G R S P D Y K -

GCGTCGTTCTCGGCACGCTGGGGGCCAACGCCGACTTCTACGAGCCCTTCGCGGACAAC  
33181 -----+-----+-----+-----+-----+ 33240  
CGCAGCAAGGAGCCGTGCGACCCCGGTTGCGGCTGAAGATGCTCGGAAGCGCTGTTG  
4 A S F L G T L G A N A D F Y E P F A D N -

GCCCGGCGCTGGTACCGGGAGTCGCAGGAGAAGGTGCTGTACTGGAACCATGCCTTCCTT  
33241 -----+-----+-----+-----+-----+ 33300  
CGGGCCGCGACCATGGCCCTCAGCGTCTCTTCCACGACATGACCTTGGTACGGAAGGAA  
4 A R R W Y R E S Q E K V L Y W N H A F L -

CACCCGCGGTCGACCGCTCGCTGCCCGCCGACGAGGTGGGCGACGCTTTCATCCACGTC  
33301 -----+-----+-----+-----+-----+ 33360  
GTGGGCGGCCAGCTGGCGAGCGACGGGCGGCTGCTCCACCGCTGCGAGAAGTAGGTGCAG  
4 H P P V D R S L P A D E V G D V F I H V -

GAGCGGGAGACCGACGCGGGCCTGGTGGTGAGCGGGGCCAAGGTGCTCGCGACCGGATCG  
33361 -----+-----+-----+-----+-----+ 33420  
CTCGCCCTCTGGCTGCGCCCGGACCACCACTCGCCCGGTTCCAGCAGCGCTGGCCTAGC  
4 E R E T D A G L V V S G A K V V A T G S -

GCCCTCACCCACGCGCGTTCATCTCGCACTGGGGACTTCCCATCAAGGACCGGAAGTTC  
33421 -----+-----+-----+-----+-----+ 33480  
CGGGAGTGGGTGCGCCGCAAGTAGAGCGTGACCCCTGAAGGGTAGTTCTTGGCTTCAAG  
4 A L T H A A F I S H W G L P I K D R K F -

GCCCTGGTGGCCACCGTGCCGATGGACGCGGACGGCCTCAAGGTGATCTGCCGTCCCTCC







35881 -----+-----+-----+-----+-----+ 35940  
 18-< GCCGATGGCGACGGTCCCGCCGAGCCCTCGCCTCCACCAGCTCAGCCACGACCAGTATA  
 A A V A A L A A D P A S T T S D T S T M -  
 CGCGGTTCCCGTCCGTTGGTTGGCGGTTTCGGCACGGCCCGAGCCCTGCCCCAGCCCCGA  
 35941 -----+-----+-----+-----+-----+ 36000  
 GCGCCAAGGGCAGGCAACCAACCGCCAAAGCCGTGCCGGGCGTCGGGACGGGCTCGGGCT  
 CGCTGGCAGGCGGCCCCGTATCAGGCATCTCCTGCGTTGCGCCCCACGCCAGTCACTTC  
 36001 -----+-----+-----+-----+-----+ 36060  
 GCGACCGTCCGCCGGGGCAGTAGTCCGTAGAGGACGCAACGCGGGGTGCGGTCACTGAAG  
 ACGGCCAGAACAAGTCGCGCATTTCTGGAAGAAGCTGAGGCCCGCGACCCGGTGGCAGCAT  
 36061 -----+-----+-----+-----+-----+ 36120  
 TGCCGGTCTTGTTTCAGCGCGTAAGACCTTCTTCGACTCCGGGCGCTGGGCCACGCTGCTA  
 CTGCGGTGTACGGAGTTTCGCACACGTTTACGCACGGAGGCTCGATGCCCCGCTGTCAATG  
 36121 -----+-----+-----+-----+-----+ 36180  
 5-> GACGCCACAGTGCCTCAAGCGTGTGCAAATGCGTGCCTCCGAGCTACGGGCGACAGTTAC  
 M P A V N G -  
 GATCGGTGCAGTCAGGCCAGTCGCACCGACGCTCCGTCGTGGCGACGGTGGTGGGCAACT  
 36181 -----+-----+-----+-----+-----+ 36240  
 5 CTAGCCACGTCACTCCGGTCAGCGTGGCTGCGAGGCAGCACCGCTGCCACCACCCGTTGA  
 S V Q S G Q S H R R S V V A T V V G N F -  
 TCGTGGAGTCGTTGACTGGCTCGCCTACGGGCTCTTCGCTCCTCTCTTCGCGGCTCAGT  
 36241 -----+-----+-----+-----+-----+ 36300  
 5 AGCACCTCAGCAAGCTGACCGAGCGGATGCCCCGAGAAGCGAGGAGAGAAGCGCCGAGTCA  
 V E S F D W L A Y G L F A P L F A A Q F -  
 TCTTCCCCCTCGTCCAACAGTTACCTCCCTGCTCGGCGCGTTCGCGGTCTTCGGCACGG  
 36301 -----+-----+-----+-----+-----+ 36360  
 5 AGAAGGGGAGCAGGTTGGTCAAGTGGAGGGACGAGCCGCGCAAGCGCCAGAAGCCGTGCC  
 F P S S N Q F T S L L G A F A V F G T G -  
 GCATGCTCTTCCGGCCGATCGGCGGGGTCTGCTGGGCGGCCCTCGCCGACCGGCGCGGCC  
 36361 -----+-----+-----+-----+-----+ 36420  
 5 CGTACGAGAAGGCCGGCTAGCCGCCCCAGGACGACCCGGCGGAGCGGCTGGCCGCGCCGG  
 M L F R P I G G V L L G R L A D R R G R -  
 GCGCCCCCGCCCTGATGCTGGCGATCGGACTGATGACCGGCGGCTCGACCCTGATCGCCG  
 36421 -----+-----+-----+-----+-----+ 36480  
 5 CCGCGGGGCGGGACTACGACCGCTAGCCTGACTACTGGCCGCGGAGCTGGGACTAGCGGC  
 R P A L M L A I G L M T G G S T L I A V -  
 TCGTCCCCACCTACGAGCACATCGGGATCCTCGCCCCGCTGCTTCTGCTGCTCGCCCCGC  
 36481 -----+-----+-----+-----+-----+ 36540  
 5 AGCAGGGGTGGATGCTCGTGTAGCCCTAGGAGCGGGGCGACGAAGACGACGAGCGGGCCG  
 V P T Y E H I G I L A P L L L L L A R L -  
 TCGCCCAGGGAGTCTCCTCGGGCGGGGAATGGACAGCGGCGGCCACCTACCTGATGGAGA  
 36541 -----+-----+-----+-----+-----+ 36600  
 5 AGCGGGTCCCTCAGAGGAGCCCCGCCCTTACCTGTGCGCCGCGGTGGATGGACTACCTCT  
 A Q G V S S G G E W T A A A T Y L M E I -  
 TCGCGCCGAAGAACC GCCGGTGCCTCTACAGCAGCCTCTTCTCCGTGACGACCATGGCGG  
 36601 -----+-----+-----+-----+-----+ 36660  
 5 AGCGCGGCTTCTTGGCGGCCACGGAGATGTGCTCGGAGAAGAGGCACTGCTGGTACCGCC  
 A P K N R R C L Y S S L F S V T T M A G -  
 GCCCCTTCGTGCGATCGCTGCTGGGCGCGGGCCTCGGCGTGTGGCTGGGAACCGCGACGA  
 36661 -----+-----+-----+-----+-----+ 36720  
 5 CGGGGAAGCAGCGTAGCGACGACCCGCGCCCGGAGCCGCACACCGACCCTTGGCGTGTCT  
 P F V A S L L G A G L G V W L G T A T M -

```

      TGGAGGCCTGGGGCTGGCGGGTGCCGTTCTCCTCGGCGGCGTCTTCGGCGTGATCCTGC
36721 -----+-----+-----+-----+-----+-----+ 36780
      ACCTCCGGACCCCGACCGCCACGGCAAGGAGGAGCCGCCGAGAAGCCGCACTAGGACG
5      E A W G W R V P F L L G G V F G V I L L -

      TGTTCTCGCGCGTCTGGCTCACCGAGACCGAGGTCTTCGCCGGGAGGTGCGGCCCCGGG
36781 -----+-----+-----+-----+-----+-----+ 36840
      ACAAGGACGCGGCAGCCGAGTGGCTCTGGCTCCAGAAGGCGGCCCTCCACGCCGGGGCCC
5      F L R R R L T E T E V F R R E V R P R A -

      CCCGGCGCGGCTCACTGGGCCAGCTGATCGGAGCCACCGCCCCAGGTGCTGCTGGCCG
36841 -----+-----+-----+-----+-----+-----+ 36900
      GGGCCGCGCCGAGTGACCCGGTCTGACTAGCCTCGGGTGGCGGGGTCCACGACGACCGGC
5      R R G S L G Q L I G A H R P Q V L L A V -

      TGATGCTGGTGGCCGGAAGTGGGCGTATCGGCGGAACGTGGTTCGACCGCGGTCCCGGCGA
36901 -----+-----+-----+-----+-----+-----+ 36960
      ACTACGACACCGGCCTGACCCGAGTAGCCGCCTTGACACAGCTGGCGCCAGGGCCGCT
5      M L V A G L G V I G G T W S T A V P A M -

      TGGGCCACCGTCTGATCGGCTCGCAGACGATGTTCTGGGTGGTGGTCTGTGTGACCGGCT
36961 -----+-----+-----+-----+-----+-----+ 37020
      ACCCGGTGGCAGACTAGCCGAGCGTCTGCTACAAGACCCACACAGACACACTGGCCGA
5      G H R L I G S Q T M F W V V V C V T G S -

      CGGTTCATCTGCTGCAGGTACCCATAGGGCTGCTCGCCGACCGGGTGAACCGGGCAGGT
37021 -----+-----+-----+-----+-----+-----+ 37080
      GCCAGTAGGACGACGTCCATGGGTATCCCGACGAGCGGCTGGCCACCTTGCCCCGTCCA
5      V I L L Q V P I G L L A D R V E P G R F -

      TCCTGATCGTCTCCAGCGTCTGCTTCGCCGCTGTGGGCTCGTACGCCTACCTACCGTCC
37081 -----+-----+-----+-----+-----+-----+ 37140
      AGGACTAGCAGAGGTTCGAGCAGAAGCGGCGACACCCGAGCATGCGGATGGAGTGGCAGG
5      L I V S S V V F A A V G S Y A Y L T V Q -

      AGGACTCCTTCGCGAGCCTGGCGTTTACGTACAGCACCGGAGTGATCTTCCTCGGCTGCG
37141 -----+-----+-----+-----+-----+-----+ 37200
      TCCTGAGGAAGCGCTCGGACCGCAAGTGCATGTCTGTGGCCTCACTAGAAGGAGCCGACGC
5      D S F A S L A F T Y S T G V I F L G C V -

      TCACCATGGTGTGCTGCCGAAGATGCTCTCCAGAATCTTCCCTCCGAGATACGCGGCCTGG
37201 -----+-----+-----+-----+-----+-----+ 37260
      AGTGGTACCACGACGGCTTCTACGAGAGGTCTTAGAAGGGAGGCGTCTATGCGCCGGACC
5      T M V L P K M L S R I F P P Q I R G L G -

      GCATCGGGCTGCCGCACGCCTCGACCACCGCACTCCTCGGCGGGGCGGGGCCACTGCTGG
37261 -----+-----+-----+-----+-----+-----+ 37320
      CGTAGCCCGACGGCGTGGGAGCTGGTGGCGTGAGGAGCCGCCCGCCCGGTGACGACC
5      I G L P H A S T T A L L G G A G P L L A -

      CCGCCTACTCCGACGAGCGAGGCGCCTCGGGCTGGTTCATCGCCGCCGTGATGGCCGCGG
37321 -----+-----+-----+-----+-----+-----+ 37380
      GCGGATGAGGCTGCTCGCTCCGCGGAGCCCGACCAAGTAGCGGCGGCACTACCGGCGCC
5      A Y S D E R G A S G W F I A A V M A A V -

      TCCTGCTCGCCTGGCCGGCCACCCTGTGGGAGCGACGGCTGTTCGCGCCCGGACGGCCC
37381 -----+-----+-----+-----+-----+-----+ 37440
      AGGACGAGCGGACCGGCCGGTGGGACACCCTCGCTGCCGACAAGGCGCGGGCCTGCCGGG
5      L L A W P A T L W E R R L F R A R T A P -

      CGGGAAGCGAGCCGGTTCCTGAATCCGCCGTCGCCCCCCCCGTCGGGTGACCGTCCGCAC
37441 -----+-----+-----+-----+-----+-----+ 37500
      GCCCTTCGCTCGGCCAAGGGCTTAGGCGGCAGCGGGCGGGGCGAGCCCACTGGCAGGCGTG
5-*      G S E P V P E S A V A R P V G * -

```

```

37501 TTCTGCATCCCGTCCGGCACCGAGCGCCGGCGACCTTCCCGACTGAGAGGTTGACATCAT
-----+-----+-----+-----+-----+ 37560
AAGACGTAGGGCAGGCCGTGGCTCGCGGCCGCTGGAAGGGCTGACTCTCCAAGTGTAGTA
23-> M -

GACGACGTCCGACACCACCGACCGGTCCCAGGACGGCGTGCCGCCGCTCTCCTTCCACCA
37561 -----+-----+-----+-----+-----+ 37620
CTGCTGCAGGCTGTGGTGGCTGGCCAGGGTCTGCGCACGGCGGCGAGAGGAAGGTGGT
23 T T S D T T D R S Q D G V P P L S F H Q -

GGAGTTCCTGTGCATGTTTCGACAGCGGGAACGACGGCGCCGACGTGGGGCCGTTTCGGCCCC
37621 -----+-----+-----+-----+-----+ 37680
CCTCAAGGACACGTACAAGCTGTGCGCCTTGCTGCGCGGGCTGCACCCCGCAAGCCGGG
23 E F L C M F D S G N D G A D V G P F G P -

CATGTACCACATCGTCGGAGCCTGGCGGCTGACCGGCGGGATCGACGAGGAGACCTTGCG
37681 -----+-----+-----+-----+-----+ 37740
GTACATGGTGTAGCAGCCTCGGACCGCCGACTGGCCGCCCTAGCTGCTCCTCTGGGACGC
23 M Y H I V G A W R L T G G I D E E T L R -

CGAGGCGCTGGGTGACGTCGTCGTGCGCCACGAGGCCCTGCGCACATCGTGGTCCGCGA
37741 -----+-----+-----+-----+-----+ 37800
GCTCCGCGACCCACTGCAGCAGCACGCGGTGCTCCGGGACGCGTGTAGCGACCAGGCGCT
23 E A L G D V V V R H E A L R T S L V R E -

AGGTGGCACGCACCGGCCGGAGATCCTGCCTGCGGGGCCCGCCGCGCTGGAGGTCCGTGA
37801 -----+-----+-----+-----+-----+ 37860
TCCACCGTGCCTGGCCGGCCTCTAGGACGCGACGCCCGGGCGGCGGCGACCTCCAGGCACT
23 G G T H R P E I L P A G P A A L E V R D -

TCTCGGCGACGTCGACGAGTCGGAGCGGGTGCGGCGCGGTGAGGAACTGCTCAACGAGGT
37861 -----+-----+-----+-----+-----+ 37920
AGAGCCGCTGCAGCTGCTCAGCCTCGCCCCACGCCGCGCCACTCCTTGACGAGTTGCTCCA
23 L G D V D E S E R V R R G E E L L N E V -

GGAGTCGACCGGTCTGAGCGTGCGGGAGCTGCCCCCTGCTGCGGGCCGTGCTCGGACGCTT
37921 -----+-----+-----+-----+-----+ 37980
CCTCAGCTGGCCAGACTCGCACGCCCTCGACGGGGACGACGCCCGGCACGAGCTTGCAGAA
23 E S T G L S V R E L P L L R A V L G R F -

CGACCAGAAGGACGCGGTGCTGGTCTCATCGCCACACACCGCCGCGGACGCTTGGGC
37981 -----+-----+-----+-----+-----+ 38040
GCTGGTCTTCTGCGCCACGACCAGGAGTAGCGGGTGGTGTGGCGGCGCCTGCGGACCCG
23 D Q K D A V L V L I A H H T A A D A W A -

CATGCACGTCATCGCCCGGACCTGCTCAACCTGTACGCCGCCAGGCGCGGGAACCCGGT
38041 -----+-----+-----+-----+-----+ 38100
GTACGTGCAGTAGCGGGCGCTGGACGAGTTGGACATGCGGCGGTCCGCGCCCTTGGGCCA
23 M H V I A R D L L N L Y A A R R G N P V -

TCCCCCGCTCCCCGAGCCGGGCCAGCATGCCGAGTTCGCCCGCTGGGAGCGCGAGGCGGC
38101 -----+-----+-----+-----+-----+ 38160
AGGGGGCGAGGGGCTCGGCCGGGTGCTACGGCTCAAGCGGGCGACCCTCGCGCTCCGCCG
23 P P L P E P A Q H A E F A R W E R E A A -

CGAGGCACCGCGGGTCGCGGTCTCGAAGGAATTCTGGCGCAAGCGCCTCCAGGGCGCGCG
38161 -----+-----+-----+-----+-----+ 38220
GCTCCGTGGCGCCAGCGCCAGAGCTTCTCTTAAGACCGCGTTTCGCGGAGGTCCCGCGCGC
23 E A P R V A V S K E F W R K R L Q G A R -

GATCATCGGGCTGGAGACGGACATACCGCGCTCGGCGGGGCTGCCCAAGGGCACCGCGTG
38221 -----+-----+-----+-----+-----+ 38280
CTAGTAGCCCGACCTCTGCCTGTATGGCGCGAGCCCGCCCGACGGTTCCCGTGGCGCAC
23 I I G L E T D I P R S A G L P K G T A W -

GCAGCGCTTCGCCGTACGCGGGAACTGGCCGACGCCGTGGTGGAGTTCTCACGGGCCGC

```





26 TGCAGGCCACTCCTTCCGCGCTGCAAGTAGCTCGAGCCTCCCGTCAGCCAGAGCCGGCAC  
 T S G E E G A T F I E L G G Q S V S A V -  
 39121 CGCATCGCCACGCGTATCCAGGAGGAGCTCGACATCTGGGTCGACATCGGCGTCTCTTTC  
 -----+-----+-----+-----+-----+ 39180  
 26 GCGTAGCGGTGCGCATAGGTCCTCCTCGAGCTGTAGACCCAGCTGTAGCCGAGGAGAAG  
 R I A T R I Q E E L D I W V D I G V L F -  
 39181 GACGACCCGGATCTGCCTACCTTCATCGCGGCGGTCTCGGACGGCCGACGCCGCGGGC  
 -----+-----+-----+-----+-----+ 39240  
 26 CTGCTGGGCCTAGACGGATGGAAGTAGCGCCGCCAGCAGGCCTGCCGGCTGCGGCGCCCG  
 D D P D L P T F I A A V V R T A D A A G -  
 39241 GCGGAGGGCTCCGGAACGCAGTGAGACTCGCCGGGCGCCGTCTCCCCGCGGCGCCCGGTT  
 -----+-----+-----+-----+-----+ 39300  
 26-\* CCGCTCCCGAGGCCTTGCCTCACTCTGAGCGGCCCGCGGCAGAGGGGCGCCGCGGGCCAA  
 G E G S G T Q \* -  
 39301 TCACATGGCTGAGGCGGTTACCCGGTACCGGTGAACCGCCTCAGCCATGTGAAACCGG  
 -----+-----+-----+-----+-----+ 39360  
 AGTGTACCGACTCCGCCAAGTGGGCCATGGCCCACTTGGCGGAGTCGGTACACTTTGGCC  
 39361 GCCTGGTCAGCGCAGCTGGATGTCCGTCTCCCGGCGATCGCCCGAGGAACTCGCCGCG  
 -----+-----+-----+-----+-----+ 39420  
 24-\* CGGACCAGTCGCGTCGACCTACAGGCAGAGGGCCCGCTAGCGGGCCTCCTTGAGCGGCGC  
 \* R L Q I D T E R A I A R L F E G R -  
 39421 GGACAGCGCGTCGGCGACCAGCTCGATGTCTCGGCCATGTACCGGTGACGCCCAGCGT  
 -----+-----+-----+-----+-----+ 39480  
 24 CCTGTGCGCGCAGCCGCTGGTTCGAGCTACAGCAGCCGGTACATGGCCAGCTGCGGGTTCGCA  
 S L A D A V L E I D D A M Y R D V G L T -  
 39481 CGGAACCAGCCGGCGCACCGCTTCGTACGTGGCCTTCGCCGCCGGGCTCAAGCCGTCGAA  
 -----+-----+-----+-----+-----+ 39540  
 24 GCCTTGGTTCGGCCGCGTGGCGAAGCATGCACCGGAAGCGGCGGCCGAGTTTCGCGAGCTT  
 P V L R R V A E Y T A K A A P S L G D F -  
 39541 CCGGCCGAGATGTGACCGCCTGGGCGGCGGCCAGGTACTCCACCGCAGGATCTTGTT  
 -----+-----+-----+-----+-----+ 39600  
 24 GGCCGGCCTCTACAGCTGGCGGACCCGCCCGGTCCATGAGGTGGCGCTCCTAGAACAA  
 R G S I D V A Q A A A L Y E V A L I K N -  
 39601 GTTGTTCGACAGGACCCGGCGGGCGTTGCGGGCCGAGATCAGGCCCATGCTCACCACGTC  
 -----+-----+-----+-----+-----+ 39660  
 24 CAACAAGCTGTCTGGGCGCCCGCAACGCCCGGTCTAGTCCGGGTACGAGTGGTGCAG  
 N N S L V R R A N R A S I L G M S V V D -  
 39661 CTGGTTGTGCGCGTTGGACGGGACGCTCTGGGTGCTGGCCGGGCGGATCGTCCGGTTCTC  
 -----+-----+-----+-----+-----+ 39720  
 24 GACCAACAGCGGCAACCTGCCCTGCGAGACCCACGACCGGCCCGGCTAGCAGGCCAAGAG  
 Q N D G N S P V S Q T S A P G I T R N E -  
 39721 GGCCACCAGTGCAGTGGCCGGGTACTGGGCGCCGGCGAATCCGCTGTGCAGCCCCGGGTC  
 -----+-----+-----+-----+-----+ 39780  
 24 CCGGTGGTCACGCCACCGGCCCATGACCCGCGGCGCTTAGGCGACACGTCGGGGCCCAG  
 A V L A T A P Y Q A G A F G S H L G P D -  
 39781 CCCGGAGACGAGGAACTCCGGGAGGCCGTAGCTGAGGTGCCGGTTCAGGACCCGGTTGAT  
 -----+-----+-----+-----+-----+ 39840  
 24 GGGCCTCTGCTCCTTGAGGCCCTCCGGCATCGACTCCACGGCCAAGTCTGGGCCAACTA  
 G S V L F E P L G Y S L H R N L V R N I -  
 39841 CTGCCGCTCGGCCAGGACGCCGAGCTGGGTGAGCGCGATGGTCACGAAGTCCATCGCGAA  
 -----+-----+-----+-----+-----+ 39900  
 GACGGCGAGCCGGTCTGCGGCTCGACCCACTCGCGCTACCAGTGCTTCAGGTAGCGCTT

24 Q R E A L V G L Q T L A I T V F D M A F -  
CGCGATCGGCTGACCGTGAAGTTTCGCCCGTGAAGATCTCCTTGCCCTCGAAGAAGAG  
39901 -----+-----+-----+-----+-----+ 39960  
GCGCTAGCCGACTGGCACCTTCAAGCGGGCACCTTCTAGAGGAACGGGAGCTTCTTCTC  
A I P Q G H F N A G H F I E K G E F F L -  
CGGGTTGTCGTTGGCCGAGTTGAGCTCGATGCGCAGCTTGTGCCGCGCGTGGTACAAGGT  
39961 -----+-----+-----+-----+-----+ 40020  
GCCCAACAGCAACCGGCTCAACTCGAGCTACGCGTGAACACGGCGCGCACCATTGTTCCA  
P N D N A S N L E I R L K H R A H Y L T -  
GTCGCGCACCGCCCCGACGACCTGGGGGATGGCCCGCAGCGAGTAGGCCTTCTGCAGGTA  
40021 -----+-----+-----+-----+-----+ 40080  
CAGCGCGTGGCGGGGCTGCTGGACCCCCCTACCGGGCGTTCGCTCATCCGGAAGACGTCCAT  
D R V A G V V Q P I A R L S Y A K Q L Y -  
GATCTCCGAGCGCTGGACGTCCTTGCCGGCCTCCTTGCTCTTCTGGAGTTCTCGGCGCAG  
40081 -----+-----+-----+-----+-----+ 40140  
CTAGAGGCTCGCGACCTGCAGGAACGGCCGGAGGAACAGGAAGACCTCAAGAGCCGCGCT  
I E S R Q V D K G A E K D K Q L E R R L -  
GTCGGCGTGTCTCGACCGTCAGTCCGCTGCCCCGCATCAGGGCCCCGCATGTTGGCGGCGGT  
40141 -----+-----+-----+-----+-----+ 40200  
CAGCCGCACGAGCTGGCAGTCAGGCGACGGGGCGTAGTCCCGGGCGTACAACCGCCGCCA  
D A H E V T L G S G R M L A R M N A A T -  
GTCGATCTGGCCCTCGTGCGGGCGGGCTATGTCTGCCCCCTCCGCGAGGAAGGGGCTGGT  
40201 -----+-----+-----+-----+-----+ 40260  
CAGCTAGACCGGGAGCACGCCCGCCGATACAGCACGGGGAGGCGCTCCTTCCCCGACCA  
D I Q G E H P R A I D H G E A L F P S T -  
CGATCCGCGTACCGCCTCGATGAGCAGAGCCGTACGATCTCGGCCTGCTGGGCCTGCTC  
40261 -----+-----+-----+-----+-----+ 40320  
GCTAGGCGCATGGCGGAGCTACTCGTCTCGGCAGTGCTAGAGCCGGACGACCCGGACGAG  
S G R V A E I L L A T V I E A Q Q A Q E -  
CAGGGCCCGTCCGACGACCAGGGAGCCAGACCGGTCTATCCCGGACGTGCCGTTGATCAG  
40321 -----+-----+-----+-----+-----+ 40380  
GTCCCGGGCAGGCTGCTGGTCCCTCGGGTCTGGCCAGTAGGGCCTGCACGGCAACTAGTC  
L A R G V V L S G L G T M G S T G N I L -  
TGCGAGGCCCTCCTTGAAGCGCAGTTTCGAGCGGCTCGATGCCCCGCTCGGCCAGCACCTG  
40381 -----+-----+-----+-----+-----+ 40440  
ACGCTCCGGGAGGAACTTCGCGTCAAGCTCGCCGAGCTACGGGGCGAGCCGGTCGTGGAC  
A L G E K F R L E L P E I G R E A L V Q -  
GGCGGTCTCCACCGGCCGTCCGTCGCGCAGGACGTAGCCCTCTCCGATGAGGGTGCTCGC  
40441 -----+-----+-----+-----+-----+ 40500  
CCGCCAGAGGTGGCCGGCAGGCAGCGCTCTGCATCGGGAGAGGCTACTCCACGAGCG  
A T E V P R G D R L V Y G E G I L T S A -  
GACGTGGGAGAGGGGAGCCAGGTCGCCGCTCGCCCCGAGTGACCCGATCTCGGGTATGGC  
40501 -----+-----+-----+-----+-----+ 40560  
CTGCACCCTCTCCCCTCGGTCCAGCGGCGAGCGGGGCTCACTGGGCTAGAGCCCATAACG  
V H S L P A L D G S A G L S G I E P I A -  
CGGGGTGATGCCCTCGTTTCAGGTACTGCGCGAGGCGTTCGAGGATGATGGGGCGCACCGC  
40561 -----+-----+-----+-----+-----+ 40620  
GCCCCACTACGGGAGCAAGTCCATGACGCGCTCCGCAAGCTCCTACTACCCCGCGTGGCG  
P T I G E N L Y Q A L R E L I I P R V A -  
GGAGTGGCCCTTGGCGAGGGTGTTTCAGCCGGGCGGCGACGATCGCCCGCGCCTCGTCTCTC  
40621 -----+-----+-----+-----+-----+ 40680  
CCTCACCGGGAACCGCTCCCAAGTCGGCCCGCGCTGTAGCGGGCGGAGCAGGAG

005070" 0070460

24 S H G K A L T N L R A A V I A R A E D E -  
GGCGAACAGCGGACCGACTCCCGCGCTGTGGCTACGGACGAGATTGGTCTGCAGTTCGAC  
40681 -----+-----+-----+-----+-----+ 40740  
CCGCTTGTGCGCTGGCTGAGGGCGCGACACCGATGCCTGCTCTAACCAGACGTCAAGCTG  
24 A F L P G V G A S H S R V L N T Q L E V -  
TTCCTTCGACTTGTGCGACCTGCATGTAGATCATCTCGCCGTACCCGGTGGTCACCCCGTA  
40741 -----+-----+-----+-----+-----+ 40800  
AAGGAAGCTGAACAGCTGGACGTACATCTAGTAGAGCGGCATGGGGCCACAGTGGGGCAT  
24 E K S K D V Q M Y I M E G Y G T T V G Y -  
GATGGGGATGTTCTGTTGCGCGATCCCTTCGAAGATCTCCCGGCTCTTCTGGGCGCTTCGC  
40801 -----+-----+-----+-----+-----+ 40860  
CTACCCCTACAAGACAAGCCGCTAGGGAAGCTTCTAGAGGGCCGAGAAGACCCGGAAGCG  
24 I P I N Q E A I G E F I E R S K Q A K A -  
GATGGATTGCGCCGGTACGTCGACCGTCGCGCGTTCCTCCGCGACGCGGCGTACGGCTTC  
40861 -----+-----+-----+-----+-----+ 40920  
CTACCTAAGCCGGCCATGCAGCTGGCAGCGCGCAAGGAGGCGCTGCGCCGCATGCCGAAG  
24 I S E A P V D V T A R E E A V R R V A E -  
GACGGTCAGGGTCTCGCCGTCGACGGAACCCGGGACGATCTCGGTCTCGACTTGAGTCAA  
40921 -----+-----+-----+-----+-----+ 40980  
CTGCCAGTCCCAGAGCGGCAGCTGCCTTTGGCCCTGCTAGAGCCAGAGCTGAACTCAGTT  
24 V T L T E G D V S V P V I E T E V Q T L -  
TGCCATCACTCCATGGGTAGCGGCCGAGGCCGGTGTACGACAGGTCAGGGGGTGGGTTCG  
40981 -----+-----+-----+-----+-----+ 41040  
ACGGTAGTGAGGTACCCATCGCCGGCTCCGGCCACATGCTGTCCAGTCCCCCACCCAAGC  
24-< A M -  
TGAGGCGCGGCTCAGCGGGTGAGCCGGGAGCGGTCCACCTTCCCCGCGGCGTTGCGCGGC  
41041 -----+-----+-----+-----+-----+ 41100  
ACTCCGCGCCGAGTCGCCCCTCGGCCCTCGCCAGGTGGAAGGGGCGCCGAACGCGCCG  
25-\* \* R T L R S R D V K G A A N R P -  
AGGCGTGAAGTCAGGCGGGTGAAGACGGCGGGCAGTGCAGGGGGGCCGAACGGCCGCGC  
41101 -----+-----+-----+-----+-----+ 41160  
TCCGCACTTCAGTCCGCCCACTTCTGCGCCCGTCACGCTCCCCGGCTTGACCGGCGCG  
25 L R S T L R T F V A P L A L P G F Q G R -  
AGATGGGAACGCCAGGCCCGGATGTCCGCGCGCACGTCTCTCCGGCCCTCTCTTGTGGC  
41161 -----+-----+-----+-----+-----+ 41220  
TCTACCCTTGGGTCCGGGCTACAGGCGCGCGTGCAGGAGGGCCGGAGAGGAACACCG  
25 L H S R W A R I D A R V D E R G E G Q P -  
ACCACGTACACGGCGAGGCGGGTCAACAGGCCCTGGCCGTTGACGTGGGGGAGGACCGCG  
41221 -----+-----+-----+-----+-----+ 41280  
TGGTGCATGTGCCGCTCCGCCAGTGGTCCGGGACCGGCAACTGCACCCCTCTTGGCGC  
25 V V Y V A L R T V L G Q G N V H P L V A -  
CACTCCAGGACCGAGGGTCAAGGTTGAGCGCGGCCTCGATCTCGGTGAGTTCCAAGCGG  
41281 -----+-----+-----+-----+-----+ 41340  
GTGAGGTCCTGGCTCCCCAGTGCCAAGTCGCGCCGAGCTAGAGCCACTCAAGGTTCCGC  
25 C E L V S P D R N L A A E I E T L E L R -  
TTCCCGAACAGCTTGACCTGGAAGTCTTGCGGCCCCGGAATTCCAGGGCTCCGTCGAAC  
41341 -----+-----+-----+-----+-----+ 41400  
AAGGGCTTGTGCAACTGGACCTTCAGGAACGCCGGGGCCTTAAGGTCCCAGGCAGCTTG  
25 N G F L K V Q F D K R G R F E L A G D F -  
CGTACCCGCGCCAGATCCCCGGTCCGGTACACCGGTACCGTCCGGGGCGAGGCCGGCG  
41401 -----+-----+-----+-----+-----+ 41460  
GCATGGGCGCGGTCTAGGGGCCAGGCCATGGTGGCCAGTGGCAGGCCCGCTCCGGCCGC  
25 R V R A L D G T R Y W R D G D P A L G A -

```

AGGGGCGCGAACAGCGCGCTGTGGTCCGGGCCGCCCTCGACGGCGAGATAACCCGGCGTC
41461 -----+-----+-----+-----+-----+-----+ 41520
TCCCCGCGCTTGTGCGCGACACCAGGCCCGGGAGCTGCCGCTCTATTGGGCCGAG
25 L P A F L A S H D P G G E V A L Y G P T -

ACGTACGGGGAGCGGATCACCAAGTTCGCCGGTGACGCCGGCGGGGCTCGGCCGGTCGTCC
41521 -----+-----+-----+-----+-----+-----+ 41580
TGCATGCCCCCTCGCCTAGTGGTCAAGCGGCCACTGCGGCCGCCCGAGCCGGCCAGCAGG
25 V Y P S R I V L E G T V G A P S P R D D -

GCGTCCACGACGAGTACCTGGCGGCCGGGGAGCGGGTACCCGATCGGGGGCCGGGCCCGTG
41581 -----+-----+-----+-----+-----+-----+ 41640
CGCAGGTGCTGCTCATGGACCGCCGGCCCCCTCGCCCATGGGCTAGCCCCGGCCCGGGCAC
25 A D V V L V Q R G P L P Y G I P A P G T -

ACCGGCCCGGTGATCTCGTGCCAGGTGCGGGCGATCGTCTCGGTGGGCCCCGTAGAGGTTG
41641 -----+-----+-----+-----+-----+-----+ 41700
TGGCCGGGCCACTAGAGCACGGTCCAGCGCCGCTAGCAGAGCCACCCGGGCATCTCCAAC
25 V P G T I E H W T A A I T E T P G Y L N -

ATCAGGCGGGTCCGGGGCAGGGCCGCGCGCAGTCCGTCCACGAGTTCGCCGGGCAGCGCC
41701 -----+-----+-----+-----+-----+-----+ 41760
TAGTCCGCCCAGGCCCCGTCCCGGCGCGCTCAGGCAGGTGCTCAAGCGGCCCGTTCGCGG
25 I L R T R P L A A R L G D V L E G P L A -

TCGCCCATCAGGAGCAGGTGGCCCCAGGGTGCCGGGCCGATCGCCCGGGTCGGAGGCGGTG
41761 -----+-----+-----+-----+-----+-----+ 41820
AGCGGGTAGTCCTCGTCCACCGGGTCCCACGGCCCCGGCTAGCGGGCCCAGCCTCCGCCAC
25 E G M L L L H G L T G P R D G P D S A T -

ATCACTCCCAGGAGGTCCCGGGCGAAGCTGGGCACGGTCTGGAGATGAGTGATCCGCTCC
41821 -----+-----+-----+-----+-----+-----+ 41880
TAGTGAGGGTCTCTCCAGGGCCCGCTTCGACCCGTGCCAGACCTCTACTCACTAGGCGAGG
25 I V G L L D R A F S P V T Q L H T I R E -

TGGACGAGCCACGGCACCAGCTTGTCGGGGTTACCCCTGACGCGCTCCGGCACCGGACAC
41881 -----+-----+-----+-----+-----+-----+ 41940
ACCTGCTCGGTGCCGTGGTTCGAACAGCCCCAAGTGGGACTGCGCGAGGCCGTGGCCTGTG
25 Q V L W P V L K D P N V R V R E P V P C -

AGCGTCCCGCCGGCCACGAGCGTCGCGAAGACCTCGGCCAGCGCCGGGTCTGTCTCCGGG
41941 -----+-----+-----+-----+-----+-----+ 42000
TCGCAGGGCGGGCCGGTGCTCGCAGCGCTTCTGGAGCCGGTTCGCGGCCAGCACGAGGCC
25 L T G G A V L T A F V E A L A P D H E P -

```

**SEQ ID No. 2. C-1027 gene cluster DNA sequence from 41,980 to 63,164**

```

AGCGCCGGGTCTGTGCTCCGGGGAGACCACTGCGCCACCCGCGCGCCCGGCCCATCGCG
41980 -----+-----+-----+-----+-----+-----+ 42039
TCGCGGCCAGCACGAGGCCCTCTGGGTGACGCGGTGGGCGCGCGGGCCGGGTAGCGC
25 L A P D H E P S V W Q A V R A G P G M A -

AACCGTTCGCCCATCCAGCCCGGAAGTGGCCCAGCGCGGCATGCGACTGGGCGATCCCC
42040 -----+-----+-----+-----+-----+-----+ 42099
TTGGCAAGCGGTAGGTTCGGGCGCTTGACCGGGTTCGCGCCGTACGCTGACCCGCTAGGGG
25 F R E G M W G A F Q G L A A H S Q A I G -

TTGGGCCCGCCCGGTTCGAACCCGAGGTGAACGCCACGTAGGCCAGGTCTGCCAGGCCCGGC
42100 -----+-----+-----+-----+-----+-----+ 42159
AACCCGGCGGGCCAGCTTGGGCTCCACTTGCGGTGCATCCGGTCCAGACGGTCCGGGCCG

```





	CGTGGAAAGTTGGCTTGCTGCTTGGCCCTGGCTCATCATCGATGGCGCTGGGCCCGCTG	
43960	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44019
	GCACCTTTTCACCGACCAGGACCGGGCGGACCGAGTAGTAGCGCTACCGCGACCCGGGCGAC	
27	R G K W L V L A A W L I I A M A L G P L -	
	GCGGGGAAGCTCGCCGACGTCCAGGACTCCAGCGCCAACGCCTTCCTTCCGCGCAGCTCG	
44020	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44079
	CGCCCCCTTCGAGCGGCTGCAGGTCCTGAGGTCGCGGTTGCGGAAGGAAGGCGCGTCGAGC	
27	A G K L A D V Q D S S A N A F L P R S S -	
	GAGTCCGCGAAGCTGAACAAGGAACTGGAGAAGTTCCGCGCCGACGAGCTGATGCCGGCC	
44080	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44139
	CTCAGGCGCTTCGACTTGTTCCTTGACCTCTTCAAGGCGCGGCTGCTCGACTACGGCCGG	
27	E S A K L N K E L E K F R A D E L M P A -	
	GTGGTGGTCTACAGCGCCGACGGCTCGCTGCCCGCCGAGGGGCGGGCCAAGGCCGAGAAG	
44140	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44199
	CACCACCAGATGTTCGCGGCTGCCGAGCGACGGGCGGCTCCCCGCCCGGTTCCGGCTCTTC	
27	V V V Y S A D G S L P A E G R A K A E K -	
	GACATAGCCGCTTCCAGGAGCTGGCCGCCGAGGGCGAGAAGGTCGAAGCGCCCCTGGAG	
44200	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44259
	CTGTATCGGCGGAAGGTCCTCGACCGGCGGCTCCCGCTCTTCCAGCTTCGCGGGGACCTC	
27	D I A A F Q E L A A E G E K V E A P L E -	
	TCCGAGGACGGCCAGGCGCTCATGGTCGTCGTTCCGCTGATCAGCGACGCCGACATCGTC	
44260	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44319
	AGCCTCCTGCGGCTCCGCGAGTACCAGCAGCAAGGCGACTAGTCGCTGCGGCTGTAGCAG	
27	S E D G Q A L M V V V P L I S D A D I V -	
	GCCACGACGAAGAAGGTCCGCGATGTCGCGGACGCCAACGCCCCCCCCGGGCGTCGCCATC	
44320	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44379
	CGGTGCTGCTTCTTCCAGGCGCTACAGCGCCTGCGGTTGCGGGGGGGCCCCGAGCGGTAG	
27	A T T K K V R D V A D A N A P P G V A I -	
	GAGGTGGGCGGGCCCCGCCGGGTCGACGACCGACGCCGCCGGCGCTTTTCGAGTCCCTCGAC	
44380	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44439
	CTCCACCCGCCCCGGGCGGCCAGCTGCTGGCTGCGGCGGCGCGAAAGCTCAGGGAGCTG	
27	E V G G P A G S T T D A A G A F E S L D -	
	TCCATGCTGATGATGGTCACCGGCCTTGTGGTCGCCATCCTGCTGCTGATCACCTACCGC	
44440	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44499
	AGGTACGACTACTACAGTGGCCGGAACACCAGCGGTAGGACGACGACTAGTGGATGGCG	
27	S M L M M V T G L V V A I L L L I T Y R -	
	TCCCCATCCTGTGGCTGCTGCCCCCTGCTCTCCGTGCGGCTTCGCCTCCGTGCTGACCCAG	
44500	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44559
	AGGGGGTAGGACACCGACGACGGGGACGAGAGGCAGCCGAAGCGGAGGCACGACTGGGTC	
27	S P I L W L L P L L S V G F A S V L T Q -	
	GTCGGCACCTACATGCTCGCCAAGTACGCCGGGCTGCCGGTCGACCCGAGAGCTCCGGC	
44560	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44619
	CAGCCGTGGATGTACGAGCGGTTTCATGCGGCCCCGACGGCCAGCTGGGCGTCTCGAGGCCG	
27	V G T Y M L A K Y A G L P V D P Q S S G -	
	GTCCTGATGGTCCTCGTGTTCCGGTGTCGGCACCGACTACGCCCTGCTGCTCATCGCCCGC	
44620	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44679
	CAGGACTACCAGGAGCACAAGCCACAGCCGTGGCTGATGCGGGACGACGAGTAGCGGGCG	
27	V L M V L V F G V G T D Y A L L L I A R -	
	TACCGTGAGGAACTGCGCCGCGAGCAGGACCGGCACGTGGCCATGAAGACCGGTTGCGA	
44680	+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	44739
	ATGGCACTCCTTGACGCGGCGCTCGTCTGCGCGTGCACCGGTACTTCTGGCGCAACGCT	
27	Y R E E L R R E Q D R H V A M K T A L R -	
	CGGTGCGGGCCCCGCCATCCTGGCCTCGGCGGGCACCATCGCCATCGGCCTCGTCTGCCTG	



**Figure 1.** The effect of the number of iterations ( $n$ ) on the accuracy of the proposed algorithm. The figure shows two plots side-by-side. The left plot shows the error norm  $\|e\|_2$  versus  $n$ , and the right plot shows the relative error  $\frac{\|e\|_2}{\|y\|_2}$  versus  $n$ . Both plots show a decreasing trend as  $n$  increases from 0 to 1000. The left plot has a logarithmic y-axis ranging from  $10^{-16}$  to  $10^0$ . The right plot has a linear y-axis ranging from 0 to 0.0008. Both plots have an x-axis ranging from 0 to 1000.

The first part of the paper discusses the importance of the
  $\mathcal{H}_2$  norm in the context of control systems. It is
 shown that the  $\mathcal{H}_2$  norm is a natural choice for
 measuring the energy of a system's response. The second
 part of the paper presents a new algorithm for computing
 the  $\mathcal{H}_2$  norm of a system. This algorithm is based on
 the use of the singular value decomposition (SVD) of the
 system's transfer function. The third part of the paper
 discusses the application of the new algorithm to the
 design of control systems. It is shown that the new
 algorithm can be used to design controllers that
 minimize the  $\mathcal{H}_2$  norm of the system's response.

[illegible]

CTCCCGGATCATCCAGATGTCCGCGGTCTGCACCGCCGGAGCCCCGGGCCATCGCGGCG  
47140 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47199  
28 GAAGGCGCTAGTAGGTCTACAGGCGCCAGACTGGCGGGCCTCGGGGCCGGGTAGCGCCGC  
S A I I Q M S A V \* (ORF28)

CGGGCCACGGCAGACAAGGAGAGAGCGTATGGCCGGCCTGGTCATGTGCGCCGGTGAGGC  
47200 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47259  
GCCCGGTGCCGTCTGTTCCTCTCTCGCATACCGGCCGACCAGTACAGCGGCCACCTCCG  
(ORF29) M A G L V M S P V E A -

GCTCGACGCGCTGGGCACGGTGCAGGGGCGTCAGGACCCCTATCCCTTCTACGAGGCGAT  
47260 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47319  
29 CGAGCTGCGCGACCCGTGCCACGTCCCCGCAGTCTTGGGGATAGGAAGATGCTCCGCTA  
L D A L G T V Q G R Q D P Y P F Y E A I -

CCGCGCGCACGGGCAGGCGGTCCCCACGAAGCCC GGCCGCTTCTGTGGTGGTTCGGCCACGA  
47320 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47379  
29 GGC GCGCGT GCCCGTCCGCCAGGGGTGCTTCGGGCCGGCGAAGCACCACCAGCCGGTGCT  
R A H G Q A V P T K P G R F V V V G H D -

CGCGTGCGACCGGGCGCTGCGGGAACCGGCCCTGCGCGTCCAGGACGCCAGGAGCTACGA  
47380 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47439  
29 GCGCACGCTGGCCCGGACGCCCTTGGCCGGGACGCGCAGGTCTGCGGTCTCTCGATGCT  
A C D R A L R E P A L R V Q D A R S Y D -

CGTCGTCTTCCCCCTCGTGGCGGTTCGCACTCCTCGGTCCGGGGGTTTACCAGCTCCATGCT  
47440 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47499  
29 GCAGCAGAAGGGGAGCACCGCCAGCGTGAGGAGCCAGGCCCCCAAGTGGTTCGAGGTACGA  
V V F P S W R S H S S V R G F T S S M L -

CTACAGCAAACCCGCCCGATCACGGCCGGTTGCGCCAGGTGGTGAGCTTCGCGTTTACCCC  
47500 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47559  
29 GATGTCTGTTGGGCGGGCTAGTGC CGGCCAACGCGGTCCACC ACTCGAAGCGCAAGTGGGG  
Y S N P P D H G R L R Q V V S F A F T P -

GCCCAAGGTGCGCCGGATGCACGGGGTGATCGAGGACATGACCGACCGGCTCCTCGACCG  
47560 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47619  
29 CGGGTTCCACGCGGCCTACGTGCCCCACTAGCTCCTGTACTGGCTGGCCGAGGAGCTGGC  
P K V R R M H G V I E D M T D R L L D R -

GATGGCCCCGGCTCGGCTCCGGCGGCTCCCCGGTCGACCTCATAGCCGAGTTTCGCCGCCCG  
47620 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47679  
29 CTACCGGGCCGAGCCGAGGCCCGGAGGGGCCAGCTGGAGTATCGGCTCAAGCGGCGGGC  
M A R L G S G G S P V D L I A E F A A R -

GCTGCCCCGTGCGGTTGATCAGCGAGATGATCGGCTTTCCGGCGAAGGACCAGGTGTGGTT  
47680 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47739  
29 CGACGGGCAGCGCCACTAGTCGCTCTACTAGCCGAAAGGCCGCTTCTGGTCCACACCAA  
L P V A V I S E M I G F P A K D Q V W F -

CCGCGACATGGCCTCCCGGGTCGCCGTGGCGACGGACGGTTTACCGACCCCGGCGCGCT  
47740 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47799  
29 GGCGCTGTACCGGAGGGCCCGAGCGGACCGCTGCCTGCCAAAGTGGCTGGGGCCGCGCGA  
R D M A S R V A V A T D G F T D P G A L -

CACGGGGGCCGACGCCGCCATGGACGAGATGAGCGCCTACTTCGACGACCTCCTGGACCG  
47800 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47859  
29 GTGCCCCCGGCTGCGGCGGTACCTGCTCTACTCGCGGATGAAGCTGCTGGAGGACCTGGC  
T G A D A A M D E M S A Y F D D L L D R -

TCGCCGCCGCACCCCGGCCGACGACCTGGTCACCCTGCTCGCCGAGGCCACGACGGCTC  
47860 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 47919  
29 AGCGGCGGCGTGGGGCCGGCTGCTGGACCAAGTGGGACGAGCGGCTCCGGGTGCTGCCGAG  
R R R T P A D D L V T L L A E A H D G S -

CCCCGGGCGCTGGACCACGACGAACTGATGGGCACCATGATGGTGCTGCTCACAGCCGG

```

47920 +-----+-----+-----+-----+-----+-----+-----+-----+ 47979
      GGGGCCCCGCGGACCTGGTGCTGCTTGACTACCCGTGGTACTACCACGACGAGTGTTCGGCC
29      P G R L D H D E L M G T M M V L L T A G -

      GTTCGAGACCACGAGCTTTCTGATCGGCCACGGGGCGATGATCGCCCTCGAACAACGGGC
47980 +-----+-----+-----+-----+-----+-----+-----+-----+ 48039
      CAAGCTCTGGTGCTCGAAAGACTAGCCGGTGCCCCGCTACTAGCGGGAGCTTGTTCGGCCG
29      F E T T S F L I G H G A M I A L E Q R A -

      GCACGCGGCCCCGGCTGCGGGCCGAACCCGACTTCGCGGACGGCTACGTTCGAGGAGATCCT
48040 +-----+-----+-----+-----+-----+-----+-----+-----+ 48099
      CGTGCGCCGGGCGGACGCCCCGGCTTGGGCTGAAGCGGCTGCCGATGCAGCTCCTCTAGGA
29      H A A R L R A E P D F A D G Y V E E I L -

      CAGGTTTCGAGCCGCGGTCCACGTCAACAGCCGGTGGGCTGCCGAGGACCTCGACCTGCT
48100 +-----+-----+-----+-----+-----+-----+-----+-----+ 48159
      GTCCAAGCTCGGCGGCCAGGTGCAGTGGTTCGGCCACCCGACGGCTCCTGGAGCTGGACGA
29      R F E P P V H V T S R W A A E D L D L L -

      GGGCCTGTCCGTACCCGCGGGCTCCAAGCTGGTCTGATCCTGGCCGCGCGGAATCGCGA
48160 +-----+-----+-----+-----+-----+-----+-----+-----+ 48219
      CCCGGACAGGCATGGCCGCCCCGAGGTTTCGACCAGGACTAGGACCGGCGGCGCTTAGCGCT
29      G L S V P A G S K L V L I L A A A N R D -

      TCCCGGCCGCTACCCCGAGCCCGGCGCTTCGACCCCGACCGCTACGCGCCCCGGCCGGG
48220 +-----+-----+-----+-----+-----+-----+-----+-----+ 48279
      AGGGCCGGCGATGGGGCTCGGGCCGGCGAAGCTGGGGCTGGCGATGCGCGGGGCGGCCCC
29      P G R Y P E P G R F D P D R Y A P R P G -

      CGGGCCGGAGGCCACCAGACCGCTGAGCTTCGGCGCGGGCGGCCACTTCTGCCTCGGCGC
48280 +-----+-----+-----+-----+-----+-----+-----+-----+ 48339
      GCCCGGCCTCCGGTGGTCTGGCGACTCGAAGCCGCGCCCGCGGTGAAGACGGAGCCGCG
29      G P E A T R P L S F G A G G H F C L G A -

      TCCGCTGGCGCGGCTGGAAGCCCGGATCGCGCTGCCGCGTCTGCTGCGCCGCTTCCCGGA
48340 +-----+-----+-----+-----+-----+-----+-----+-----+ 48399
      AGGCGACCGCGCCGACCTTCGGGCCTAGCGCGACGGCGCAGACGCGCGCGAAGGGCCT
29      P L A R L E A R I A L P R L L R R F P D -

      CCTGGCCGTGTCCGAGCCCCCGTCTACCGCGACCGCTGGGTTCGTCCGCGGCCTCGAAAC
48400 +-----+-----+-----+-----+-----+-----+-----+-----+ 48459
      GGACCGGCACAGGCTCGGGGGCAGATGGCGCTGGCGACCCAGCAGGCGCCGGAGCTTTG
29      L A V S E P P V Y R D R W V V R G L E T -

      CTTTCCCGTGACCCTCGGGTCTGAGCCCCCGCGGCCGGAACACGTGACCGTCCCGGCC
48460 +-----+-----+-----+-----+-----+-----+-----+-----+ 48519
      GAAAGGGCACTGGGAGCCCAGGACTCGGGGGCGGCCGCTTGTGCACTGGCAGGGCCGG
29      F P V T L G S * (ORF29)

      GCGGGGTGCGCGCCCTCTCAGACGTACAGGGTGTGGGGCCCTGACCACACAGCACCCGG
48520 +-----+-----+-----+-----+-----+-----+-----+-----+ 48579
      CCGCCACGCGCGGGAGAGTCTGCATGTCCACAAACCCGGGGACTGGTGTGTCTGTTGGCC

      CCGTACAGCTCCAGGTTGGTGCTCGGGTTCATGCAGGTGCAGCGTGATGCTCTGGGCATC
48580 +-----+-----+-----+-----+-----+-----+-----+-----+ 48639
      GGCATGTGCGAGGTCCAACCACGAGCCCCAGTACGTCCACGTGCACTACGAGACCCGTAG
30      (ORF30)* A P A A H H E P C

      GCTGCACGCGCTGGATCGGGACGTCTGTGTAGATCGAGGACCCGCGCTCGCCTGGGCGA
48640 +-----+-----+-----+-----+-----+-----+-----+-----+ 48699
      CGACGTGCGCGACCTAGCCCTGCAGCAACACTAGCTCCTGGGCGGCGAGCGGACCCGCT
30      R Q V R Q I P V D N Y I S S G G S A Q A -

      GGATGTCCACCGACTCCTTGCCCAGTCGGCACGCCCCGCCCCAGCAGGCCGCGGCACAGCA
48700 +-----+-----+-----+-----+-----+-----+-----+-----+ 48759
      CCTACAGGTGGCTGAGGAACGGGTTCAGCCGTGCGGGCGGGTTCGTCCGGCGCCGTGTCGT

```





50440 GTGGCGCGGGTGTACCGGGAGACCGAGGAGCAGTTCGGCGTGCTCGCGCCCCCTGGCC 50499  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CACCGCGCCACATGGCCCTCTGGCTCCTCGTCAAGCCGACGAGCGCGGGGGGACCGG  
 V A R V Y R E T E E Q F G V L A P P L A -  
  
 50500 CTCCACTCGCCCCCGCGGGCGTCTGGCCGCGACGTGGCTCATGCTGCGGGAGACACTG 50559  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 GAGGTGAGCGGGCGCGCCGACGCGACCGGCGCTGCACCGAGTACGACGCCCTCTGTGAC  
 L H S P A A A S L A A T W L M L R E T L -  
  
 50560 CTGGTCGACGGGCGGGTGAGCCGGGCGGTGAAGGAGACGGTCGCCACCGAGGTCTCCCGT 50619  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 GACCAGTGCCTCGCCCACTCGGCCCCCACTTCCTCTGCCAGCGGTGGCTCCAGAGGGCA  
 L V D G R V S R A V K E T V A T E V S R -  
  
 50620 GCCAACGACTGTCCGTACTGCGTCCAGGTCCATCAGGCGGTACTCGGGACACTGCCTCCG 50679  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CGGTTGCTGACAGGCATGACGCGAGGTCCAGGTAGTCCGCCATGAGCCCTGTGACGGAGGC  
 A N D C P Y C V Q V H Q A V L G T L P P -  
  
 50680 GACGGCGGCCAGGCCGGGCTCCTGCGGTGGGTCCGGGAGGCAGGCCGACGGCCCGGGCGGC 50739  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CTGCCGCGGTCCGGCCCCGAGGACGCCACCCAGGCCCTCCGTCCGGCTGCCGGGCCGCGC  
 D G G Q A G L L R W V R E A G R R P G G  
  
 50740 GGTGCGGTGGGCGGCGGGCGGCCGCTTCCGTTTCCGCGGTGAACAGGCACCGGAAGTGTGC 50799  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CCACGCCACCCGCGCCCGCGGCGAAGGCAAGTCGCCCACTTGTCCGTGGCCTTGACACG  
 G A V G G G R P L P F S G E Q A P E L C -  
  
 50800 GGCCTCGTGGTCACGTTCCACTACATCAACCGCATGGTCTCCCTCTTCTCGACGACTCC 50859  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CCGCAGCACCAGTGCAAGGTGATGTAGTTGGCGTACCAGAGGGAGAAGGAGCTGCTGAGG  
 G V V V T F H Y I N R M V S L F L D D S -  
  
 50860 CCCATGCCGACCCGACGCCGACACCGTTGCGCGGGGCCCATCATGAGGACCACCGCACTG 50919  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 GGGTACGGCTGGGCTGCGGCTGTGGCAACGCGCCCGGTAGTACTCCTGGTGGCGTGAC  
 P M P T R T P T P L R G P I M R T T A L -  
  
 50920 GCCATGCGTCCCGTCGGCCCCGGGCTGCTGACACCGGGCGCATCGCTCGGCCTGCTGCCT 50979  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CGGTACGACAGGGCAGCCGGGCCCCGACGACTGTGGCCCGCGTAGCGAGCCGGACGACGGA  
 A M R P V G P G L L T P G A S L G L L P -  
  
 50980 CCGGCTCCCCTGCCGCCCGGACTGGAGTGGGCGGAGGGCAACCCTTTCTGTTGGCCAGGCG 51039  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 GGCCGAGGGGACGGCGGGCCTGACCTCACCCGGCTCCCGTTGGGAAAGCACCGGTTCCGC  
 P A P L P P G L E W A E G N P F V A Q A -  
  
 51040 CTGGGGCGTGCCGTCGCCGCTGTGGACCAGGGAGCGCACTGGGTGCCCCGAACCGGTCCGG 51099  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 GACCCCGCACGGCAGCGGCGACACCTGGTCCCTCGCGTGACCCACGGGCTTGGCCAGGCC  
 L G R A V A A V D Q G A H W V P E P V R -  
  
 51100 GAGCGGCTGCGCACACGTCTGGACACCTGGGACGGATCGGCGCCGGGCTCGGCCGGGGA 51159  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 CTCGCCGACGCGTGTGACAGCCTGTGGACCCTGCCTAGCCGCGGCCCGGAGCCGGCCCCCT  
 E R L R T R L D T W D G S A P G L G R G -  
  
 51160 TGGCTCGACGAGGCCGTGTCCGGCCTGCCGCCCCAGGACGTGCCCCGCGGCACGGCTGGCG 51219  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+  
 31 ACCGAGCTGCTCCGGCACAGGCCGGACGGCGGGGTCTGACGGGCGCGTGGCCGACCGC  
 W L D E A V S G L P P Q D V P A A R L A -  
  
 CTGCTGACGGCCTTCGCCCCCTACCAGGTGCTCCCGGACGACGTGAGGAGTTGACACGG



	51220	+-----+-----+-----+-----+-----+-----+	51279
		GACGACTGCCCGGAAGCGGGGATGGTCCACGAGGCCTGCTGCAGCTCCTCAAGTCTGCC	
31		L L T A F A P Y Q V L P D D V E E F R R -	
		CGTCGGCCCCACCGACCGCGAACTCGTTCGAGCTCACGTCTACGCCGCGCTGACCACGGCC	
	51280	+-----+-----+-----+-----+-----+-----+	51339
		GCAGCCGGGTGGCTGGCGCTTGAGCAGCTCGAGTGCAGGATGCGGCGCGACTGGTGCCGG	
31		R R P T D R E L V E L T S Y A A L T T A -	
		GTCCGTGTCCGGTCGCACGCTCGTTCGTGCCCCACGCCGCCGGGCGGGGATGAACGGCCCCG	
	51340	+-----+-----+-----+-----+-----+-----+	51399
		CAGGCACAGCCAGCGTGCAGCAGCAGCGGCTGCGGCGGCCCGGCCCTACTTGCCGGGGC	
31		V R V G R T L V V P D A A G P G * (ORF31)	
		CAACGGCTCGGGAAGGCTGTCTCACGGCCGGAGGCGTACGCCGGTGAGGTGCTCGGACTC	
	51400	+-----+-----+-----+-----+-----+-----+	51459
		GTTGCCGAGCCCTTCCGACAGAGTGCCGGCCTCCGCATGCGGCCACTCCACGAGCCTGAG	
		(ORF32) * P R L R V G T L H E S E -	
		CTCCCAGAGGCGGCGCCGGGCCCTGGGGTCGACGGCTGCTCCGCCGGGGCGCACGAGCCC	
	51460	+-----+-----+-----+-----+-----+-----+	51519
		GAGGGTCTCCGCGCGGCCCGGGACCCAGCTGCCGACGAGCGCGGCCCGCGTGCTCGGG	
32		E W L R R R A R P D V A A G G P R V L G -	
		GGGTGCGCCCCGGGTCTCGGTACGCCGAGGGGCCCGTAGAACTCGCCCCGCGCGCGCC	
	51520	+-----+-----+-----+-----+-----+-----+	51579
		CCCACGCGGGGCCAGAGCCAGTGCGGCTCCCGGGCATCTTGAGCGGGGCGCGCGCGG	
32		P A G R T E T V G L P G Y F E G G R A G -	
		GGGATCGGTGGCCGCCCGCAGACCAGGCAGCATCCCCGCCGCGCGGGCTGCAGGAACAA	
	51580	+-----+-----+-----+-----+-----+-----+	51639
		CCCTAGCCACCGGCGGGCGTCTGGTCCGTCTAGGGCGGCGCCGCCGACGTCCTTGTT	
32		P D T A A R L G P L M G A A A P Q L F L -	
		CGGGGCGAGCGGGGAGCCGAGCCTGCGCACGGGCGCGGAAAGTCCCGGCCAGACCGGT	
	51640	+-----+-----+-----+-----+-----+-----+	51699
		GCCCCGCTCGCCCCCTCGGCTCGGACGCGTGCCCCGCGCCCTTTCAGGGCCGGGTCTGGCCA	
32		P A L P S G L R R V P A P F D R G L G T -	
		CGCGGTACGCCCCGGGATGAGCGGCGAGCGAGGCCAGTTCGCGCCGGA CTCCGCCAGTCT	
	51700	+-----+-----+-----+-----+-----+-----+	51759
		GCGCCAGTCGGGCCCTACTCGCCGCTCGCTCCGGTCAAGGCGCGGCCTGAGGCGGT CAGA	
32		A T L G P H A A L S A L E A G S E A L R -	
		GTGATGGAGTTCCAGCGCGAACATGAGGTTGGCCAGCTTGGACTGGTTGTAGGCCCGGTA	
	51760	+-----+-----+-----+-----+-----+-----+	51819
		CACTACCTCAAGGTCGCGCTTGTA CTCCAACCGT CGAACCTGACCAACATCCGGGCCAT	
32		H H L E L A F M L N A L K S Q N Y A R Y -	
		CCGGCTGTAGCGGCGTTCCGCCGTGAAGGTCGCTGAAGTCGATGCGCCCCAGCCGGTGCAG	
	51820	+-----+-----+-----+-----+-----+-----+	51879
		GGCCGACATCGCCGCAAGCGGC ACTTCAGCGACTTCAGCTACGCGGGGTGCGCCACGTC	
32		R S Y R R E G H L D S F D I R G L R H L -	
		ATAGCTGCTGATCGTTCACGACCCGCGCGCCCGGCGCGGCCCGCAGGCTGTCCAGGAGCAG	
	51880	+-----+-----+-----+-----+-----+-----+	51939
		TATCGACGACTAGCAGTGCTGGGCGCGCGGGCCGCGCCGGGCGTCCGACAGGT CCTCGTC	
32		Y S S I T V V R A G P A A R L S D L L L -	
		GCCGGTGAGGGCGAAGTGCCCCAGGTGGTTCTGTGGCGAACTGGAGTTCGTGACCGTCCGG	
	51940	+-----+-----+-----+-----+-----+-----+	51999
		CGGCCACTCCCGCTTCACGGGGTCCACCAAGCACCGCTTGACCTCAAGCACTGGCAGGCC	
32		G T L A F H G L H N T A F Q L E H G D P -	
		GGTGCGGGCCCCGGTCCGATCACATCACGCCCGCGTTGTTGACCAGCAGGTGGATGCGCGG	
	52000	+-----+-----+-----+-----+-----+-----+	52059

```

32      CCACGCCCCGGGCCAGCCAGGTGTAGTGCGGGCGCAACAACCTGGTCGTCCACCTACGCGCC
      T R A R D T W M V G A N N V L L H I R P -

      GAAGCGGTTCGCGCAGTTCTCTCGGCGCCGGCACGCACCGACGCGAGACGGGAAAGATCCAG
52060 +-----+-----+-----+-----+-----+-----+-----+----- 52119
      CTTTCGCCAGCGCGTCAAGGAGCCGCGGCCGTGCGTGGCTGCGCTCTGCCCTTTCTAGGTC
32      F R D R L E E A G A R V S A L R S L D L -

      CCGTCTGACCGTCAGTTGCGCCGACGGCACCCGGCTTTGGATGCGGGCCGCGCGGCGAC
52120 +-----+-----+-----+-----+-----+-----+-----+----- 52179
      GGCAGACTGGCAGTCAACGCGGCTGCCGTGGGCCGAAACCTACGCCCCGGCGGCGCCGCTG
32      R R V T L Q A S P V R S Q I R A A A A V -

      CCCGCGGTCCGGATCGCGCACGGCCAGCACACGCTGGGCGCCGTGCCGGGCGAGCTCCTG
52180 +-----+-----+-----+-----+-----+-----+-----+----- 52239
      GGGCGCCAGGCCTAGCGCGTGCCTGGTGTGTCACCCGCGGCACGGCCCCGCTCGAGGAC
32      G R D P D R V A L V V H A G H R A L E Q -

      CGCCAGGTGCAGTCCGATGCCGGAGCTGGCACCGGTGACCACCGCGGTGGTTCCGGTACG
52240 +-----+-----+-----+-----+-----+-----+-----+----- 52299
      GCGGTCCACGTCAAGCTACGGCCTCGACCGTGGCCACTGGTGGCGCCACCAAGGCCATGC
32      A L H L G I G S S A G T V V A T T G T R -

      GTCCGGGACATCGGCGGCGCTCCAGCGTGCCTCGCTTCTCATCGGTGCTCCCTCCCGGGG
52300 +-----+-----+-----+-----+-----+-----+-----+----- 52359
      CAGGCCCTGTAGCCGCGCGAGGTGCGAGCGGCGCAAGAGTAGCCAGCAGGAGGGCCCC
32      D P V D A A S W R R R T R M (ORF32)

      GATGCGTCAGCCGGCCTGGGCCATCGCGGCCCGGTAGCCGTTGGCGACGATCTGCCGGGC
52360 +-----+-----+-----+-----+-----+-----+-----+----- 52419
      CTACGCAGTCGGCCGGACCCGGTAGCGCCGGGCCATCGGCAACCGCTGCTAGACGGCCCC

      GGAGTGCTCGTAGTACTCGTCGTCTTCGGCAGCTCCGTGGCGAGACCGCTGACGTACCG
52420 +-----+-----+-----+-----+-----+-----+-----+----- 52479
      CCTCACGAGCATCATGAGCAGCAGGAAGCCGTGAGGCACCGCTCTGGCGACTGCATGGC

      GTTGAACATGCAGAACGCGGCGGCGATCAGAACGGTGTGTCGAGAGCGGTGTGTCGTCGCG
52480 +-----+-----+-----+-----+-----+-----+-----+----- 52539
      CAACTTGTCAGTCTTGCGCCGCCGCTAGTCTTGCCACAGCACGTCTGCCACAGCAGGCG

      TCCCTCGGCCCCGCGCCGAGGCGATCACCCCTGCGGAGACCGGGCGCGCCGCGCTCTGGAC
52540 +-----+-----+-----+-----+-----+-----+-----+----- 52599
      AGGGAGCCGGGCGCGGCTCCGCTAGTGGGGACGCCTCTGGCCCCGCGGGCGGAGACCTG

      CTCGGCGGCGACGGCCAGCAGCGCGCGCTCCTGCCGTGATGGGCGCGGTGGCGGGGTC
52600 +-----+-----+-----+-----+-----+-----+-----+----- 52659
      GAGCCGCCGCTGCCGGTCTGTCGCGCGCGCAGGACGGCAGCTACCCGCGCCACCGCCCCAG

      GGCAGGACGGCCTCGACGAGCTGCCGGCCTCCCGGCAGCTGCGCGGCGGCGAAGGCCCC
52660 +-----+-----+-----+-----+-----+-----+-----+----- 52719
      CCGCTCCTGCCGGAGCTGCTCGACGGCCGAGGGCCGTGACGCGCCGCCGCTTCCGGGG

      GTGGGAGGCGGCGCAGAACTCGGTGGAGTTGAGATGCGAGACGTACGCCGCGATGAGCTC
52720 +-----+-----+-----+-----+-----+-----+-----+----- 52779
      CACCCTCCGCCGCGTCTTGAGCCACCTCAACTCTACGCTCTGCATGCGGCGCTACTCGAG

      GCGTTGCCCCGGTTCCAGCGAGGACGGCGCCCGCAGCAGGGCGTTGCGGAGATCGCCAG
52780 +-----+-----+-----+-----+-----+-----+-----+----- 52839
      CGCAACGGGGCCAAGGTCGCTCCTGCCGCGGGCGTCTGCCGCAAGCGCTCTAGCGGGTC

      CGGTGCTGCGGTGCCGGGGTGGTGAGCCATCAGACCACTGATGCCGGGGAGGTGCTTGTC
52840 +-----+-----+-----+-----+-----+-----+-----+----- 52899
      GCCACGACGCCACGGCCCCACCACTCGGTAGTCTGGTGACTACGGCCCCCTCAGCAACAG

      GAGTGCTATGTGGGGCACGGCTCTTCTTCCGGGTGGACGAGGGGCGGACGGCGGCGGAT
52900 +-----+-----+-----+-----+-----+-----+-----+----- 52959

```

[illegible]



[illegible]



35	A H S A Q T A L R N T G R L A A A A A L -	
56320	CGCGGGAGCGAGGTCGTCTCTGCTCATGGAGGGGTCCGAGGACATCGACACCGTGTCTGGCC	56379
35	GCGCCCTCGCTCCAGCAGGACGAGTACCTCCCCAGGCTCCTGTAGCTGTGGCACGACCGG	
	R G S E V V L L M E G S E D I D T V L A -	
56380	GTCCATGACGCCCTGGTGAACCGTTACGACAACGTGGGGATCACCCCTTCAGGCGCACCTG	56439
35	CAGGTACTGCGGGACCACTTGGAATGCTGTTGCACCCCTAGTGGGAAGTCCGCGTGGAC	
	V H D A L V N R Y D N V G I T L Q A H L -	
56440	CACCGCACCGTGGACGACGCCATGGCGGTTCGCGGGTCTGGCCGCACCGTGC GGCTGGTC	56499
35	GTGGCGTGGCACCTGCTGCGGTACCGCCAGCGCCAGGACCGGCGTGGCACGCCGACCAG	
	H R T V D D A M A V A G P G R T V R L V -	
56500	ATGGGCTCCTCGGCCGAGCCTGCCGGCACCGCTCTGTCCCGGGGCCCCGCTCTGGAGGAC	56559
35	TACCCGAGGAGCCGGCTCGGACGGCCGTGGCGAGACAGGGCCCCGGGGCGAGACCTCCTG	
	M G S S A E P A G T A L S R G P A L E D -	
56560	CGGTACCTTGACCTCGCGGAGCTTCTCGTGGACCGTGGCGTCCGGCTGAGTCTGGCCACT	56619
35	GCCATGGAAGTGGAGCGCTCGAAGAGCACCTGGCACCGCAGGCCGACTCAGACCGGTGA	
	R Y L D L A E L L V D R G V R L S L A T -	
56620	CCGGACGCCGAGGTCCTGGCCGGGGCGCAGGAGCGTGGTCTGCTCGAACGCGTCCAGGAC	56679
35	GGCCTGCGGCTCCAGGACCGGCCCGCGTCTCGCACACGAGCTTGCGCAGGTCCTG	
	P D A E V L A G A Q E R G L L E R V Q D -	
56680	ATCGAGATGCTCTACGGTGTGCGGCCCGAGCTGCTGCGCCGCCACCGGGCGGCGGGCCGC	56739
35	TAGCTCTACGAGATGCCACACGCCGGGCTCGACGACGCGGCGGTGGCCCCGCCGCCGCG	
	I E M L Y G V R P E L L R R H R A A G R -	
56740	CCCTGTGCGATCCACGCGGCCTACGGGATGAACTGGTGGCTTCCCCTGCTGCGGAGGCTG	56799
35	GGGACAGCGTAGGTGCGCCGATGCCCTACTTGACCACCGAAGGGGACGACGCCTCCGAC	
	P C R I H A A Y G M N W W L P L L R R L -	
56800	GCCGACAACCCGCCGATGGTGTCTCAACGCCCTGGCCGACATCGGCCGGGACCGGGAGCCC	56859
35	CGGCTGTTGGGCGGCTACCACGAGTTGCGGGACCGGCTGTAGCCGGCCCTGGCCCTCGGG	
	A D N P P M V L N A L A D I G R D R E P -	
56860	GTCGCCCACCAGGCGTACTGACCCGCCCCGGGCCGCGATCCGCGGGGCACCGGCCCGGG	56919
35	CAGCGGGTGGTCCGCGATGACTGGGCGGGGCCCGGCGCTAGGCGCCCCGTGGCCGGGGCCC	
	V A H Q A Y * (OF35)	
56920	GCGCCGGTCAGCTCCCGGTGCGCCGCGAACTGCCCGGGCCTGCGCCCCCTCGCCCGCCGGCC	56979
36	CGCGGCCAGTCGAGGGCCAGCGGCGCTTGACGGGCCCCGACGCGGGGAGCGGGCGGCCGG	
	(ORF36) * S G T A A F Q G P R R G E G A P -	
56980	CCCGGTAGGCCTGGGCGATGTCCAGCCACTTCTCCGCCTCCTGACCAGACGCGGTACGGG	57039
36	GGGCCATCCGGAACCGCTACAGGTGCGGTGAAGAGGCGGAGGACTGGTCTGCGCCAGTCCC	
	G R Y A Q A I D L W K E A E Q G S A T L -	
57040	CGAGGTCGTGCGGTTGGCGGCGCCGGGTGACCAGCAGGCAGAAGTCGTGCGCGGGACCGC	57099
36	GCTCCAGCAGCGCCACCGCCGCGGCCCACTGGTCTCGTCCGTCTTCAGCACGCGCCCTGGCG	
	A L D D R H R R R T V L L C F D H A P G -	

57100 TGACCGTCTCGGTGGCGTCCTCGGGGCCGACCGTCCAGACCTCGCCCGAGGGGGCGGTGA 57159  
 +-----+-----+-----+-----+-----+-----+  
 36 ACTGGCAGAGCCACCGCAGGAGCCCCGGCTGGCAGGTCTGGAGCGGGCTCCCCGCCACT -  
 S V T E T A D E P G V T W V E G S P A T  
  
 57160 GCTCGAAGCGGAACGGCGCGGCCGGGGTTCAGACCGTGGGACTCGTAGCCGAAGTCGC 57219  
 +-----+-----+-----+-----+-----+-----+  
 36 CGAGCTTCGCCTTGCCGCGCCGGCCGCCAGTCTGGCACCTGAGCATCGGCTTCAGCG -  
 L E F R F P A A P P T L G H S E Y G F D  
  
 57220 GTGTCAGCCAGGCGAAGTCGACGATGTTGCGAAGCCGCTCGGTGGGCGTGCGCCGACAC 57279  
 +-----+-----+-----+-----+-----+-----+  
 36 CACAGTCGGTCCGCTTCAGCTGCTACAACGCTTCGGCGAGCCACCCGCACGCGGCTGTG -  
 R T L W A F D V I N R L R E T P T R R V  
  
 57280 CCAGGGCGTCGGCGACGTCCTGGCCGTGGGCGAACACCTCCATGATCCCGGCGCAGCCCA 57339  
 +-----+-----+-----+-----+-----+-----+  
 36 GGTCCCGCAGCCGCTGCAGGACCGGCACCCGCTTGTGGAGGTACTAGGGCCGCGTCGGGT -  
 G L A D A V D Q G H A F V E M I G A C G  
  
 57340 GAACGACCGGCGGCAGCGGGTTGACCAGCCACGGAACACCTGGCCGGCGGGGACCGCGG 57399  
 +-----+-----+-----+-----+-----+-----+  
 36 CTTGCTGGCCCGCGTCGCCCACTGGTTCGGTGCCTTGGTGGACCGGCCCGCCCTGGCGCC -  
 L V V P P L P N V L W P V V Q G A P V A  
  
 57400 CGAGCGCCTCGACCGAGGCCCCGCCCATGCCCGGAAGCGGGTGAGCAGTTCCTGCGGCG 57459  
 +-----+-----+-----+-----+-----+-----+  
 36 GCTCGCGGAGCTGGCTCCGGGCGGGGTACGGGGCCTTCGCCACTCGTCAAGGACGCCGC -  
 A L A E V S A R G M G R F R T L L E Q P  
  
 57460 GGAAGCCCTTGAAGTCTGTCAGAGCCGCGTTGACCGCTCCGTGCAAGTTGCCTGCCGCGG 57519  
 +-----+-----+-----+-----+-----+-----+  
 36 CCTTCGGGAAGTTCAGACGCTCTCGGCGCAACTGGCGAGGCAGCTTCAACGGACGGCGCC -  
 P F G K F Q Q L A A N V A G D F N G A A  
  
 57520 CGGCCGTGACGGCCTTGAAGTCTCTCCGGCGCCGCCCGCGGTCTGGCCAGGTTGAAGA 57579  
 +-----+-----+-----+-----+-----+-----+  
 36 GCCGGCACTGCCGGAAGTTCAGGAGGCCGCGGCGGCGGCCAGGACCGGTCCAAGTCTCT -  
 A A T V A K F E E P A A A A T R A L N F  
  
 57580 CGAAGGTGAGGTGGGCGATCTGGTTCGGTGACGGTCCAGCCGGGCGCCGGCGTTCGGAGTGT 57639  
 +-----+-----+-----+-----+-----+-----+  
 36 GCTTCCACTCCACCCGCTAGACCAGCCACTGCCAGGTTCGGCCCGCGGCCGAGCCTCACA -  
 V F T L H A I Q D T V T W G P A P T P T  
  
 57640 TCCAGGCTTCGTCGTCGATCTTCTCGACCAGCTGCGCCAGCTCCTCGATGTCGGTGGCCA 57699  
 +-----+-----+-----+-----+-----+-----+  
 36 AGGTCCGAAGCAGCAGCTAGAAGAGCTGGTTCGACGCGGTTCGAGGAGCTACAGCCACCGGT -  
 N W A E D D I K E V L Q A L E E I D T A  
  
 57700 GGTGCTTGAGGACGTCGTCGAGCGAATTCATCTCGTACTTCCTTCACTGGGGGTGTTCCG 57759  
 +-----+-----+-----+-----+-----+-----+  
 36 CCACGAACTCCTGCAGCAGCTCGCTTAAGTAGAGCATGAAGGAAGTGACCCCCACAAGGC -  
 L H K L V D D L S N M (ORF36)  
  
 57760 GGCTGGGACGGATGTCCCGCCGGGTGGGCGCGGCCGGCGGAAGCGCCGTCGCGGAGCG 57819  
 +-----+-----+-----+-----+-----+-----+  
 CCGACCCTGCCTACAGGGCGGCCACCCGGCCCGGCCGCTTCGCGGCAGCGCCTCGC  
  
 57820 TCGGCGACAGTCGCTAGGCGGCGGTCCCGCGTAGGAGCCGGCCCGGTTCGGAATAGGGCG 57879  
 +-----+-----+-----+-----+-----+-----+  
 37 AGCCGCTGTGACGATCCGCCGCGCAGGGCGCATCCTCGGCCGGGCCAGCCTTATCCCGC -  
 (ORF37) \* A A R G A Y S G A R D S Y P  
  
 CGAGCGCCTCGGCCAGGGCTTCGGGTATCAGGGTCGGCACGGTCGCCGTGTTGGGGCCGC



	57880	+-----+-----+-----+-----+-----+	57939
		GCTCGCGGAGCCGGTCCCGAAGCCCATAGTCCCAGCCGTGCCAGCGGCACAACCCCGGCG	
37		A L A E A L A E P I L T P V T A T N P G -	
		GCATGCAGGCGATGCGCTGGCGTCCCCGCGCCACCAGGGTCTCGCCGCCGTTCGTGCGCCA	
	57940	+-----+-----+-----+-----+-----+	57999
		CGTACGTCCGCTACGCGACCGCAGGGGCGCGGTGGTCCCAGAGCGGCGGCAGCAGCGGGT	
37		R M C A I R Q R G R A V L T E G G D D G -	
		GCTTGATGTAGTCGAAGGTGAACTCCAGCTGGGTCTGCCGCAGCTCCGAGAGCCTCATCC	
	58000	+-----+-----+-----+-----+-----+	58059
		CGAACTACATCAGCTTCCACTTGAGGTGCACCCAGACGGCGTTCGAGGCTCTCGGAGTAGG	
37		L K I Y D F T F E L Q T Q R L E S L R M -	
		GGATCGACAGTTTCGTGCAAGGCGGTGATCTCCGCGAAGAAGTTCGCAGTCCACCTTGAGGG	
	58060	+-----+-----+-----+-----+-----+	58119
		CCTAGCTGTCAAGCAGCTTCCGCCACTAGAGGCGCTTCTTGAGCGTCAGGTGGAAGTCCC	
37		R I S L E D F A T I E A F F E C D V K L -	
		TGAAGAGCTTGAGGTCTCTCGGACCTCGGCGAGCACCGAAGGCGCCCTCTCCTTGAGAA	
	58120	+-----+-----+-----+-----+-----+	58179
		ACTTCTCGAACTCCAGGAGGACCTGGAGCCGCTCGTGGCTTCCGCGGGAGAGGAAGTCTT	
37		T F L K L D E Q V E A L V S P A R E K L -	
		AGAGTTCCCCGCAACGCCCCCTGCCAACGAAGGTAGTTGACGTAGTAGACGTTGCCGACGA	
	58180	+-----+-----+-----+-----+-----+	58239
		TCTCAAGGGCCGTTGCGGGGACGGTTGCTTCCATCAACTGCATCATCTGCAACGGCTGCT	
37		F L E R C R G Q W R L Y N V Y Y V N G V -	
		GGTTCGTCTCCTCGAAGCCGACGGTGTGGCGGAGCTCGAAGTAGTCAGGATTTCGTGCGGG	
	58240	+-----+-----+-----+-----+-----+	58299
		CCAAGCAGAGGAGCTTCGGCTGCCACACCGCCTCGAGCTTCATCAGTCCTAAGCAGCGCC	
37		L N T E E F G V T H R L E F Y D P N T A -	
		TCATAGGTCTGTGCCCTTCGTGCTCGGGGCGGTCGTGCGACCGAGTTGCGTGAAGCAAC	
	58300	+-----+-----+-----+-----+-----+	58359
		AGTATCCAGACACGGAAGCAGCAGCCCCGCGCAGCAGCGTGGCTCAACGCACTTCGTTG	
37		T M (ORF37)	
		TCACTGGTTCGCGATGGCCTGCGGGGTTCGGTGGCCCGCGCTCCGGGCGGAGAGTGGGGCG	
	58360	+-----+-----+-----+-----+-----+	58419
		AGTGACCAGCGCTACCGGACGCCCCAGCCACCGGGCGCGAGGCCCGCCTCTCACGCCCCG	
		GGGTGCCGGCCGGCGCGGGGTTCAGCCGCGCGCCGACGGCAGCAGGGGAAGAACCCTCTCG	
	58420	+-----+-----+-----+-----+-----+	58479
		CCCACGGCCGGCGCGCCCCAGTCGGCGCGCGGCTGCGTTCCTTCTTGGGAGAGC	
38		(ORF38) * G R A S P L L P L V R E -	
		CGGCCGCTCGTGGAGCCGTTCGGGGGCGGTGCGCCGTAGGTGACGGAGATAACCCGGGCTC	
	58480	+-----+-----+-----+-----+-----+	58539
		GCCGCGAGCACCTCGGCAGCCCCGCGCCACGCGGCATCCACTGCCTCTATGGGGCCGAG	
38		R G S T S G D P A P A G Y T V S I G R S -	
		TGCGCGGCGCGCACGATCCCCGGCATCGCGGTTTCGGCGAGCGCCGCGATGGTCATCGCG	
	58540	+-----+-----+-----+-----+-----+	58599
		ACGCGCCGCGCGTGTAGGGGCGGTAGCGCGCAAGCCGCTCGCGGCGCTACCACTAGCGC	
38		Q A A R V I G P M A R E A L A A I T M A -	
		GGATTGACCGTCAGCGCGCCGGGAACCGACGATCCGTTCGGTGACGAAGATCCCCGGGTGG	
	58600	+-----+-----+-----+-----+-----+	58659
		CCTAACTGGCAGTCGCGCGGCCCTTGGCTGCTAGGCAGCCACTGCTTCTAGGGGCCCCACC	
38		P N V T L A G P V S S G D T V F I G P H -	
		TCGCGGAGCTCGTTGCTGTGCTCCAGGGCGGATGTGTGGGGGTTCGTGCGCCATCCGGCAG	
	58660	+-----+-----+-----+-----+-----+	58719
		AGCGCCTCGAGCAACGACAGCAGGTCCCGCTACACACCCCCAGCAGCGGGTAGGCCGTC	











[illegible]